CAN EVERYTHING BE COMPUTED? - ON THE SOLVABILITY COMPLEXITY INDEX AND TOWERS OF ALGORITHMS

J. BEN-ARTZI, A. C. HANSEN, O. NEVANLINNA, AND M. SEIDEL

ABSTRACT. This paper addresses and establishes some of the fundamental barriers in the theory of computations and finally settles the long standing computational spectral problem.

Due to the barriers presented in this paper, there are many problems, some of them at the heart of computational theory, that do not fit into the classical frameworks of complexity theory. Hence, we are in need for a new extended theory of complexity, capable of handling these new issues. Such a theory is presented in this paper. Many computational problems can be solved as follows: a sequence of approximations is created by an algorithm, and the solution to the problem is the limit of this sequence (think about computing eigenvalues of a matrix for example). However, as we demonstrate, for several basic problems in computations (computing spectra of infinite dimensional operators, solutions to linear equations or roots of polynomials using rational maps) such a procedure based on one limit is impossible. Yet, one can compute solutions to these problems, but only by using several limits. This may come as a surprise, however, this touches onto the definite boundaries of computational mathematics. To analyze this phenomenon we use the Solvability Complexity Index (SCI). The SCI is the smallest number of limits needed in order to compute a desired quantity. In several cases (spectral problems, inverse problems) we provide sharp results on the SCI, thus we establish the absolute barriers for what can be achieved computationally. For example, we show that the SCI of spectra and essential spectra of infinite matrices is equal to three, and that the SCI of spectra of self-adjoint infinite matrices is equal to two, thus providing the lower bound barriers and the first algorithms to compute such spectra in two and three limits. This finally settles the long standing computational spectral problem. We also show that the SCI of solutions to infinite linear systems is two.

Moreover, we establish barriers on error control. We prove that no algorithm can provide error control on the computational spectral problem or solutions to infinite-dimensional linear systems. In particular, one can get arbitrarily close to the solution, but never knowing when one is ”epsilon” away. This is universal for all algorithms regardless of operations allowed. In addition, we provide bounds for the SCI of spectra of classes of Schrödinger operators, thus we affirmatively answer the long standing question on whether or not these spectra can actually be computed. Finally, we show how the SCI provides a natural framework for understanding barriers in computations. It has a direct link to the Arithmetical Hierarchy, and in particular, we demonstrate how the impossibility result of McMullen on polynomial root finding with rational maps in one limit, and the framework of Doyle and McMullen on solving the quintic in several limits, can be put in the SCI framework.

CONTENTS

Date: August 14, 2015.

1. Introduction

Let us start with the old and famous computational problem: can one compute the zeros of a polynomial by using only finitely many arithmetic operations and radicals of the polynomial coefficients? The answer is of course well known: yes, if the degree is less than five, no, if the degree is five or higher. However, what if we are also allowed to pass to a limit? In particular, can one construct a sequence of sets, where the construction of each element requires finitely many arithmetic operations and radicals, such that this sequence converges to the zeros of the polynomial? Indeed, the answer is yes, and this gives us a technique to compute zeros of polynomials (through a limiting process). Given the equivalence between roots of polynomials and spectra of matrices, this also provides a way of computing eigenvalues of matrices, given the matrix values. But what happens if the matrix becomes infinite (as an operator on $l^2(\mathbb{N})$): can one then compute the spectrum by using finitely many arithmetic operations and radicals and then pass to the limit as in the finite-dimensional case?

Consider another fundamental computational problem. Given a Schrödinger operator $H = -\Delta + V$ that is defined on some appropriate domain so that the spectrum is uniquely determined by the potential, can one compute the spectrum by using arithmetic operations and radicals on the point samples $V(x)$ where $x \in \mathbb{R}^d$ and then take limits? Note that this problem differs from the matrix problem in that one can only access $V(x)$ and not matrix elements.

The former problem was only recently partially solved affirmatively in [49]. However, the computation in this case is done with three limits, not one as is needed in the finite-dimensional case. This begs the question: is this a barrier that cannot be broken, does one really need several limits to do the computation? Given that the former problem was only recently partially solved, it is not a surprise that the latter problem is still open (although solutions exists in some special cases). It is an intriguing observation that it is more than eighty years since Schrödinger won his Nobel Prize in physics, yet how to compute spectra of arbitrary Schrödinger

operators is still unknown. The key question is therefore: can this be done with a finite number of limits (as in the matrix case), and if so what is the smallest number of limits needed?

Let us consider another basic problem in computation. What if we want to compute a root of a polynomial, however we are not allowed to use radicals, but rather a rational map applied iteratively (such as Newton's method). The problem with Newton's method is that it may not converge. This problem prompted S. Smale [93] to ask whether there exists an alternative to Newton's method, namely, a purely iterative generally convergent algorithm (see Section 8 for definition). Smale asked: "*Is there any purely iterative generally convergent algorithm for polynomial zero finding ?*" His conjecture was that the answer is 'no'. This problem was settled by C. McMullen in [67] as follows: yes, if the degree is three; no, if the degree is higher (see also [68, 95]). However, in [38] P. Doyle and C. McMullen demonstrated a striking phenomenon: this problem can be solved in the case of the quartic and the quintic using several limits. In particular, they provide a construction such that, by using several rational maps and independent limits, a root of the polynomial can be computed.

It turns out that the examples above touch onto the absolute barriers in the theory of computations. In particular, many of these highly important problems cannot be solved by passing to a single limit, however, they can be solved by using several limits. This is a rather intriguing phenomenon and is important as more limits make the computations more complex. In this paper we provide a unifying framework for these kinds of questions by introducing a general concept of the Solvability Complexity Index (SCI) and towers of algorithms. This is done by generalizing the frameworks in [49] and [38]. The SCI of a computational problem is the least amount of limits needed in order to compute the desired quantity, given a certain set of allowed operations. Many of the basic computational problems (from the insolvability of the quintic and the negative answer to Smale's question, to questions on computing spectra of infinite matrices or Schrödinger operators, solutions to linear systems and even the Arithmetical Hierarchy) fit into this abstract framework. Moreover, we provide sharp bounds on the SCI for several of these problems and thus demonstrate solutions to - and barriers for - fundamental questions in computational mathematics. Some of the highlights are the following.

Main results:

(i) It is impossible to compute spectra and essential spectra of infinite matrices in less than three limits. This is universal for all algorithms regardless of the operations allowed (arithmetic operation, radicals etc). The only assumption on the algorithms is that they can only read a finite amount of information in each iteration step. This implies that even if there had existed an algorithm that could compute the spectrum of a finite dimensional matrix using finitely many arithmetic operations (of course no such algorithm exists), one could still not compute the spectrum of an infinite matrix in less than three limits. However, it is possible to compute spectra and essential spectra in three limits when allowing arithmetic operations of complex numbers. This finally settles the general computational spectral problem as considered in [49].

(ii) It is impossible to compute spectra of self-adjoint infinite matrices in less than two limits. This is, as above, universal and regardless of the operations allowed. However, it is possible to compute spectra of matrices with controlled growth on their resolvent in two limits (this then includes normal operators). It is possible to compute spectra of tridiagonal self-adjoint infinite matrices in one limit, but getting the essential spectrum is impossible, for that one needs two limits.

(iii) One can compute spectra of all non-normal Schrödinger operators with bounded potential with bounded local total variation in two limits. If the operator is normal the number of limits is one (actually we prove much more, only knowledge on the growth of the resolvent is needed). If the potential blows up at infinity, the spectrum can be computed in one limit, regardless of non-normality. This establishes an affirmative answer to the long standing problem of computing spectra of Schrödinger

operators for a vast class of potentials. Note that the techniques used here also pave the way for even more classes of potentials.

(iv) It is impossible to compute the solution to a general infinite linear system in one limit (this is universal for all algorithms regardless of the operations allowed), yet it is possible in two. The same goes for the problem of computing the norm of the inverse of an infinite matrix. For matrices with known/controllable off diagonal decay, one can compute the solution to a linear system in one limit.

(v) Many problems that have SCI greater than one can never be computed with error control. In other words it is impossible to design an algorithm that can compute an approximation to the solution and know when one is ”epsilon” away from the true solution. This is the case of many spectral problems and also the case for solutions to linear systems. Thus, in these cases, no-one can ever know with certainty, how close the computation is to the true solution. This is a universal statement for any possible algorithm.

(vi) The SCI framework provides a new complexity theory for problems that do not fit into the existing complexity theories. In particular, current complexity theory cannot handle problems that requires several limits in the computation. We predict that this class of problems is vast, although this paper only exhibit a subset of this potentially big class. However, it already contains some of the core problems in computations such as spectral problems and inverse problems as well as polynomial root finding with rational maps.

Note also that the SCI concept addresses the basic problem of computing with limits (Problem 5, p. 43, [12]) as posed by Blum, Shub and Smale.

(vii) When considering decision problems there is a clear connection between the SCI and the Arithmetical Hierarchy. In particular, the Δ_m sets in the Arithmetical Hierarchy can equivalently be characterised in term of the SCI. Thus, one may view the SCI as a classification tool that is a generalisation of this complexity hierarchy to arbitrary computational problems. The link to the Arithmetical Hierarchy implies that the SCI can become arbitrarily large, i.e. for any $k \in \mathbb{N}$ there is a problem such that the SCI is equal to k.

Remark 1.1 (Polynomial time algorithms and implementation). Note that all upper bounds on the SCI provided in this paper are proved constructively and hence yield actual implementable algorithms that can be used in practice. Moreover, all of them are indeed fast, meaning that the output produced, for each step of the iterations, is done in polynomial time in the number of input the algorithm reads. For actual implementations see [50]. See also [48, 49].

Remark 1.2 (Existing complexity theory vs. the Solvability Complexity Index). Note that, without going into details, classical complexity theory can essentially be summed up by the following two classes of problems:

(I) The problem can be solved in finite time, and the task is to analyze how difficult this is. The complexity theory of these types of problem hosts the famous “P versus NP” problem that is one of the major open problems in mathematics today [99].

(II) The problem cannot be solved in finite time, however, it can be solved by computing a sequence of approximations, and the solution to the problem is the limit of the approximations. The task is to analyze the difficulty of carrying out the approximation procedure. The complexity theory of such problems contains famous questions, among them Smale’s 17th problem [7,23], and a very rich mathematical theory [11,24,90,93,94]. The field of information based complexity theory [71,74,96] is devoted specifically to these types of problems.

The key issue is that, as this paper establishes, there are fundamental barriers that prevent many problems at the heart of computational theory from fitting into the above frameworks. In particular, there is a third class of problems.

(III) The problem cannot be computed by approximations and then passing to a limit, however, it can be computed by approximations and then passing to several limits.

The theory of the Solvability Complexity Index offers a new view of complexity theory that bridges this gap. The theory is very flexible and does indeed incorporate the classical approaches. In particular, the class I discussed above is included in the set of problems that have SCI equal to zero, and the class II is included in the set of problems with SCI equal to one. This new extended complexity theory does not have to assume a machine (although it certainly can, as discussed in Section 7.2) like the Turing machine [98] or a Blum-Shub-Smale machine [12]. The theory of the SCI and towers of algorithms only specify which mathematical operations are allowed. This allows essentially most computational problems into this framework ranging from analysis to recursion theory.

Remark 1.3 (The SCI and a new classification theory). Note that although this paper establishes several classification results with the SCI, it leaves a vast set of open problems. In particular, one must now classify all problems according to their SCI. And it is of utmost importance to get an understanding of what kind of structure and extra information that allows one to lower the index. We predict that the class of problems with SCI greater than one is vast, see Section 9 for details.

How to read the paper: For the reader only interested in the main results, Section 2 is the only prerequisite. The following sections after Section 2 present the main results from different parts of mathematics, however, these are completely self-contained. If the reader is interested in the proof techniques, an easy introduction is in the proofs of some of the basic decision problems. Thus, Section 7.1 is a good start followed by Section 14.

2. The Solvability Complexity Index and towers of algorithms

Throughout this paper we assume the following:

$$\Omega \text{ is some set, called the } \textit{primary} \text{ set,} \tag{2.1a}$$

$$\Lambda \text{ is a set of complex valued functions on } \Omega, \text{ called the } \textit{evaluation} \text{ set,} \tag{2.1b}$$

$$\mathcal{M} \text{ is a metric space,} \tag{2.1c}$$

$$\Xi \text{ is a mapping } \Omega \to \mathcal{M}, \text{ called the } \textit{problem} \text{ function.} \tag{2.1d}$$

The set Ω is essentially the set of objects that give rise to our computational problems. It can be a family of matrices (infinite or finite), a collection of polynomials, a family of Schrödinger (or Dirac) operators with a certain potential etc. The problem function $\Xi : \Omega \to \mathcal{M}$ is what we are interested in computing. It could be the set of eigenvalues of an $n \times n$ matrix, the spectrum of a Hilbert (or Banach) space operator, root(s) of a polynomial etc. Finally, the set Λ is the collection of functions that provide us with the information we are allowed to read, say matrix elements, polynomial coefficients or pointwise values of a potential function of a Schrödinger operator, for example.

In most cases it is convenient to consider a metric space $\mathcal{M}$, however, in the case of polynomials it may be more useful to use a pseudo metric space (see Example 2.1 (III)). To explain this rather abstract setup in (2.1) we commence with the following examples:

Example 2.1. (I) **(Spectral problems)** Let $\Omega = \mathcal{B}(\mathcal{H})$, the set of all bounded linear operators on a separable Hilbert space $\mathcal{H}$, and the problem function Ξ be the mapping $A \mapsto \mathrm{sp}(A)$ (the spectrum of

A). Here $(\mathcal{M}, d)$ is the set of all compact subsets of $\mathbb{C}$ provided with the Hausdorff metric $d = d_{\mathrm{H}}$ (defined precisely in (3.1)). The evaluation functions in Λ could for example consist of the family of all functions $f_{i,j} : A \mapsto \langle Ae_j, e_i\rangle$, $i, j \in \mathbb{N}$, which provide the entries of the matrix representation of A w.r.t. an orthonormal basis $\{e_i\}_{i\in\mathbb{N}}$. Of course, Ω could be a strict subset of $\mathcal{B}(\mathcal{H})$, for example the set of self-adjoint or normal operators, and Ξ could have represented the pseudo spectrum, the essential spectrum or any other interesting information about the operator.

(II) **(Inverse problems)** Let $\Omega = \mathcal{B}_{\mathrm{inv}}(\mathcal{H}) \times \mathcal{H}$, where $\mathcal{B}_{\mathrm{inv}}(\mathcal{H})$ denotes the set of all bounded invertible operators on $\mathcal{H}$, and let the problem function Ξ be the mapping $(A, b) \mapsto A^{-1}b$, which assigns to a linear problem $Ax = b$ its solution x. The metric space $\mathcal{M}$ would simply be $\mathcal{H}$ and Λ the collection of mappings $\{f_{i,j}\}_{i\in\mathbb{N},j\in\mathbb{Z}_+}$ where $f_{i,j} : (A, b) \mapsto \langle Ae_j, e_i\rangle$ for $j \in \mathbb{N}$ and $f_{i,0} : (A, b) \mapsto \langle b, e_i\rangle$. Also here Ω could consist of operators with specific properties (off diagonal decay, self-adjointness, isometric properties).

(III) **(Polynomial root finding)** Let $\Omega = \mathbb{P}_s$, the set of polynomials of degree $\leq s$ over $\mathbb{C}$ and let the problem function Ξ be the mapping $p \mapsto \{\alpha \in \mathbb{C} \mid p(\alpha) = 0\}$ (the roots of p). Let $(\mathcal{M}, d)$ denote the collection of finite sets of points in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ equipped with the pseudo metric $d : \mathcal{M} \times \mathcal{M} \to [0, \infty]$, defined by $d(x, y) = \min_{1\leq i\leq n, 1\leq j\leq m} |x_j - y_i|$, where $x = \{x_1, \dots, x_n\}, y = \{y_1, \dots, y_m\} \in \mathcal{M}$. The reason for the pseudo metric is that the techniques of Doyle and McMullen that we will consider are based on computing a single root of a polynomial (as for example Newton's method does). In this case Λ is the finite set of functions $\{f_j\}_{j=1}^s$ where $f_j : p \mapsto \alpha_j$ for $p(t) = \sum_{k=1}^s \alpha_k t^k$.

(IV) **(Computational quantum mechanics)** Let $\Omega = \mathrm{L}^\infty(\mathbb{R}^d) \cap \mathrm{C}(\mathbb{R}^d)$ and let $\Xi : V \mapsto \mathrm{sp}(-\Delta + V)$, where the domain $\mathcal{D}(-\Delta + V) = \mathrm{W}^{2,2}(\mathbb{R}^d)$ (the standard Sobolev space) and $-\Delta + V$ is the usual Schrödinger operator. Given that the spectra are unbounded, we cannot use the Hausdorff metric anymore, but will let $(\mathcal{M}, d_{\mathrm{AW}})$ denote the set of closed subsets of $\mathbb{C}$ equipped with the *Attouch-Wets* metric (see (4.2)). In this case a natural choice of Λ would be the set of all evaluations $f_x : V \mapsto V(x)$, $x \in \mathbb{R}^d$.

(V) **(Decision making)** Let Ω denote the set of infinite matrices with values in $\{0, 1\}$ and $\Xi : \Omega \to \mathcal{M} = \{\textit{Yes}, \textit{No}\}$ where $\mathcal{M}$ is equipped with the discrete metric d_{disc}. The evaluation functions would naturally be $f_{i,j} : A \mapsto A_{i,j}$, $i, j \in \mathbb{N}$, the (i, j)th matrix coordinate of A. A typical example of Ξ could be: $\Xi(\{A_{i,j}\})$: Does $\{A_{i,j}\}$ have a column containing infinitely many non-zero entries? Naturally, Ω can be replaced with the natural numbers including zero $\mathbb{Z}_+$ and Ξ could be a question about membership in a certain set, as in classical recursion theory. In this case the evaluation set would be $\Lambda = \{\lambda\}$ consisting of the function $\lambda : \mathbb{Z}_+ \to \mathbb{C}$, $x \mapsto x$.

Given this setup and motivation, we can now define what we mean by a computational problem.

Definition 2.2 (Computational problem). *Given a primary set Ω, an evaluation set Λ, a (pseudo) metric space $\mathcal{M}$ and a problem function $\Xi : \Omega \to \mathcal{M}$ we call the collection $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ a computational problem.*

Our aim is to find and to study families of functions (that we will sometimes refer to as algorithms) which permit to approximate the function Ξ. The main pillar of our framework is the concept of a tower of algorithms. However, before that we will define a general algorithm.

Definition 2.3 (General Algorithm). *Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$, a* general algorithm *is a mapping $\Gamma : \Omega \to \mathcal{M}$ such that for each $A \in \Omega$:*

(i) *there exists a finite subset of evaluations $\Lambda_\Gamma(A) \subset \Lambda$,*
(ii) *the action of Γ on A only depends on $\{A_f\}_{f\in\Lambda_\Gamma(A)}$ where $A_f := f(A)$,*
(iii) *for every $B \in \Omega$ such that $B_f = A_f$ for every $f \in \Lambda_\Gamma(A)$, it holds that $\Lambda_\Gamma(B) = \Lambda_\Gamma(A)$.*

We will sometimes write $\Gamma(\{A_f\}_{f\in\Lambda_\Gamma(A)})$, *in order to emphasize that* $\Gamma(A)$ *only depends on the results* $\{A_f\}_{f\in\Lambda_\Gamma(A)}$ *of finitely many evaluations.*

Note that for a general algorithm there are no restrictions on the operations allowed. The only restriction is that it can only take a finite amount of information, though it *is* allowed to *adaptively* choose the finite amount of information it reads depending on the input (which may very well be infinite, say an infinite matrix, or a function). The condition (iii) just ensures that the algorithm is well defined and consistent since, put in simple words, changing the input A shall not affect the algorithm's action as long as the change does not affect the output of the relevant evaluations in $\Lambda_\Gamma(A)$.

Definition 2.4 (Tower of algorithms). *Given a computational problem* $\{\Xi,\Omega,\mathcal{M},\Lambda\}$, *a* tower of algorithms of height k for $\{\Xi,\Omega,\mathcal{M},\Lambda\}$ *is a family of sequences of functions*

$$\begin{aligned}\Gamma_{n_k} &: \Omega\to\mathcal{M},\\ \Gamma_{n_k,n_{k-1}} &: \Omega\to\mathcal{M},\\ &\vdots\\ \Gamma_{n_k,\ldots,n_1} &: \Omega\to\mathcal{M},\end{aligned}$$

where $n_k,\ldots,n_1\in\mathbb{N}$ *and the functions* $\Gamma_{n_k,\ldots,n_1}$ *at the lowest level in the tower are general algorithms in the sense of Definiton 2.3. Moreover, for every* $A\in\Omega$,

$$\begin{aligned}\Xi(A) &= \lim_{n_k\to\infty}\Gamma_{n_k}(A),\\ \Gamma_{n_k}(A) &= \lim_{n_{k-1}\to\infty}\Gamma_{n_k,n_{k-1}}(A),\\ &\vdots\\ \Gamma_{n_k,\ldots,n_2}(A) &= \lim_{n_1\to\infty}\Gamma_{n_k,\ldots,n_1}(A),\end{aligned}\tag{2.2}$$

where $S=\lim_{n\to\infty}S_n$ *means convergence* $S_n\to S$ *in the (pseudo) metric space* $\mathcal{M}$.

In this paper we will discuss several types of towers: *Doyle-McMullen towers*, *Kleene-Shoenfield towers*, *Arithmetic towers*, *Radical towers* and *General towers*. A General tower will refer to the very general definition in Definition 2.4 specifying that there are no further restrictions as will be the case for the other towers. When we specify the type of tower, we specify requirements on the functions $\Gamma_{n_k,\ldots,n_1},\ldots,\Gamma_{n_1}$ in the hierarchy, in particular, what kind of operations may be allowed. Thus, a tower of algorithms for a computational problem is essentially the toolbox allowed. The Doyle-McMullen tower appeared first in the paper of Doyle and McMullen [38] (but then only referred to as a tower of algorithms). The Kleene-Shoenfield towers describe the Arithmetical Hierarchy known from Classical Recursion Theory as we will see in Section 7. A Radical tower, as defined below, first appeared in [49] where it was referred to as a "set of estimating functions" for computing spectra. The definition here is substantially more general and allows for the use of these types of towers for a wide range of problems.

Given the definition of a tower of algorithms we can now define the main concept of this paper: the Solvability Complexity Index (SCI). The SCI was first discussed in [49] for a specific spectral problem, however, this definition extends to include general problems in computations.

Definition 2.5 (Solvability Complexity Index). *Given a computational problem* $\{\Xi,\Omega,\mathcal{M},\Lambda\}$, *it is said to have* Solvability Complexity Index $\mathrm{SCI}(\Xi,\Omega,\mathcal{M},\Lambda)_\alpha=k$ *with respect to a tower of algorithms of type* α *if* k *is the smallest integer for which there exists a tower of algorithms of type* α *of height* k. *If no such tower exists then* $\mathrm{SCI}(\Xi,\Omega,\mathcal{M},\Lambda)_\alpha=\infty$. *If there exists a tower* $\{\Gamma_n\}_{n\in\mathbb{N}}$ *of type* α *and height one such that* $\Xi=\Gamma_{n_1}$ *for some* $n_1<\infty$, *then we define* $\mathrm{SCI}(\Xi,\Omega,\mathcal{M},\Lambda)_\alpha=0$.

The key goal with the SCI is that many problems can be put into this framework and analyzed. In particular, given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ and a tower of algorithms of a certain type, we are interested in the question: what is the smallest height of a tower possible? If it is finite, it means that the problem can be computed, if it is infinite it cannot be computed with such a tower. However, if it is finite, this number says something about how difficult it is to compute. Having several limits in the computation makes the computation more demanding. As we will see later in the paper many problems have SCI greater than one. Note that the SCI is a characteristic of a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$. Thus, if any of the elements $\Omega, \mathcal{M}, \Lambda$ are changed the SCI may also change. As we will see in many cases, the change of domain Ω changes the SCI, as may changes of $\mathcal{M}$ and Λ.

We can now define an *Arithmetic tower of algorithms* and a *Radical tower of algorithms* (for the definition of a *Doyle-McMullen tower of algorithms* see Definition 8.1 in Section 8, for *Kleene-Shoenfield* see Definition 7.8 in Section 7, and notice that both are special types of Arithmetic towers in these special settings).

Definition 2.6 (Arithmetic and Radical towers). *Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ we define the following:*

(i) *An* Arithmetic tower of algorithms *of height k for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ is a tower of algorithms where the lowest functions $\Gamma = \Gamma_{n_k,\dots,n_1} : \Omega \to \mathcal{M}$ satisfy the following: For each $A \in \Omega$ the action of Γ on A consists of only performing finitely many arithmetic operations on $\{A_f\}_{f\in\Lambda_\Gamma(A)}$ where we remind that $A_f = f(A)$.*

(ii) *A* Radical tower of algorithms *of height k for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ is a tower of algorithms where the lowest functions $\Gamma = \Gamma_{n_k,\dots,n_1} : \Omega \to \mathcal{M}$ satisfy the following: For each $A \in \Omega$ the action of Γ on A consists of only performing finitely many arithmetic operations on and extracting radicals of $\{A_f\}_{f\in\Lambda_\Gamma(A)}$.*

Remark 2.7. To state the definition of Arithmetic towers in other words one may say that for the finitely many steps of the computation of the lowest functions $\Gamma = \Gamma_{n_k,\dots,n_1} : \Omega \to \mathcal{M}$ only the four arithmetic operations $+, -, \cdot, /$ within the smallest (algebraic) field which is generated by the input $\{A_f\}_{f\in\Lambda_\Gamma(A)}$ are allowed. When we use the word radical we mean that we can evaluate the mapping $a \mapsto \sqrt[n]{a}$ for $a \geq 0$ and $n \in \mathbb{N}$. Also, in both Arithmetic and Radical towers we implicitly assume that any complex number can be decomposed into a real and an imaginary part, and moreover we can determine whether $a = b$ or $a > b$ for all real numbers a, b which can occur during the computations.

Remark 2.8. In this paper we will mostly be concerned with Arithmetic towers (as opposed to Radical towers used in [49]). The reason is that most of the problems considered in this paper can be solved without resorting to radicals. It should be mentioned that in a practical setting a Radical tower may be preferable for stability reasons. This is beyond the scope of this paper and we refer to [50] for details.

When considering the SCI of a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ with respect to a tower of certain type we will use the following notation:

$$\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{DM}}, \qquad \mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{KS}},$$

$$\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{A}}, \quad \mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{R}}, \quad \mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}}$$

to denote the SCI with respect to Doyle-McMullen, Kleene-Schoenfield, Arithmetic, Radical and General towers respectively. Note the inequalities: $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{A}} \geq \mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{R}} \geq \mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}}$ that are obvious. Given the rather abstract setup, let us consider some simple and clarifying examples.

Example 2.9. (I) **(Spectra of $n \times n$ matrices and roots of polynomials)** Let $\Omega_1 = \mathbb{C}^{4\times 4}$ and $\Omega_2 = \mathbb{C}^{5\times 5}$ and let $\Xi : A \mapsto \mathrm{sp}(A)$. Let also $\mathcal{M}$ and Λ be defined as in Example 2.1 (I) above. Then,

obviously $\mathrm{SCI}(\Xi, \Omega_1, \mathcal{M}, \Lambda)_{\mathrm{R}} = 0$ as it is possible to express the eigenvalues of the matrix by using finitely many arithmetic operations and radicals of the matrix elements. However, the eigenvalues cannot be expressed without radicals, thus $\mathrm{SCI}(\Xi, \Omega_1, \mathcal{M}, \Lambda)_{\mathrm{A}} > 0$. The situation is different when considering Ω_2. In particular, because of the insolvability of the quintic using radicals we must have $\mathrm{SCI}(\Xi, \Omega_2, \mathcal{M}, \Lambda)_{\mathrm{R}} > 0$. Moreover, showing the insolvability of the quintic is obviously equivalent to showing $\mathrm{SCI}(\Xi, \Omega_2, \mathcal{M}, \Lambda)_{\mathrm{R}} > 0$. Is it then clear that $\mathrm{SCI}(\Xi, \Omega_2, \mathcal{M}, \Lambda)_{\mathrm{R}} = 1$? Note that none of the standard approaches such as the QR-algorithm gives this result as they are not globally convergent. However, as a result of the developments in [49] it is relatively easy to show that $\mathrm{SCI}(\Xi, \Omega_1, \mathcal{M}, \Lambda)_{\mathrm{R}} = 1$.

(II) **(Spectra of compact operators)** Let $\Omega = \mathcal{K}(l^2(\mathbb{N})) \subset \mathcal{B}(l^2(\mathbb{N}))$ denote the set of compact operators on $l^2(\mathbb{N})$. We assume access to the matrix elements, so Λ is as above, as is $\mathcal{M}$. To build a Radical tower of height two for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ consider the following. Define P_m to be the projection onto $\mathrm{span}\{e_1, \dots, e_m\}$ (where $\{e_n\}_{n\in\mathbb{N}}$ is the canonical basis). For $A \in \Omega$, define $\Gamma_{m,n}(A) = \hat{\Gamma}_{m,n}(P_m A P_m|_{\mathrm{Ran}(P_m)})$, and $\Gamma_m(A) = \mathrm{sp}(P_m A P_m|_{\mathrm{Ran}(P_m)})$, where $\{\hat{\Gamma}_{m,n}\}_{n\in\mathbb{N}}$ is a Radical tower of algorithms such that $\hat{\Gamma}_{m,n}(B) \to \mathrm{sp}(B)$ for every $B \in \mathcal{B}(\mathrm{Ran}(P_m))$ and every m, as $n \to \infty$. Note that this is possible by the example above with $m \times m$ matrices. In particular, we have $\Gamma_m(A) = \lim_{n\to\infty} \Gamma_{m,n}(A)$ and $\Xi(A) = \lim_{m\to\infty} \Gamma_m(A)$ [39, Part II, XI.9 Lemma 5], which yields the bound $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{R}} \leq 2$. Actually, we will show in Theorem 3.7 that $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{A}} = 1$.

Example 2.10 (Assumptions on Λ). The set Λ decides what the algorithm can read, and depending on what Λ may contain, the SCI may change as the following example demonstrates:

> Given a cylinder of radius $r > 0$ and height 1, compute a cube, more precisely its edge length a, of the same volume.

Thus, $\Omega = (0, \infty)$ shall be the set of all such radii, $\mathcal{M} = (0, \infty)$, Λ simply contains the evaluation $f : r \mapsto r$, and the problem function $\Xi : \Omega \to \mathcal{M}$ maps each radius r to the desired edge length a, respectively. Formally, we have $a = \sqrt[3]{\pi r^2}$, and we immediately see that $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{A}} \geq \mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{R}} \geq 1$: otherwise there would exist an algorithm for the computation of a, only performing finitely many arithmetic operations on and extracting radicals of the radius r (and the interim results). This algorithm would particularly work on the restricted problem $(\Xi, \{1\}, \mathcal{M}, \Lambda)$. Consequently there would also be a finite Radical algorithm for the computation of $a^3 = \pi$, contradicting the fact that π is transcendental.

Now one may assume that this one critical constant π is already available (from e.g. some oracle, some precomputation or somewhere else). Then there are two questions arising: How can we provide the Arithmetic (resp. Radical) algorithms with this additional tool? and what is the resulting SCI?

For the former there is a very simple solution: Just supplement Λ by a further evaluation function, namely the constant function $g : r \mapsto \pi$ which outputs π for every input $r \in \Omega$. By doing this one actually adapts the computational problem $(\Xi, \Omega, \mathcal{M}, \Lambda)$ to the new problem $(\Xi, \Omega, \mathcal{M}, \tilde{\Lambda})$ with $\tilde{\Lambda} = \{f, g\}$: "Given r compute a, under the preassumption that π is already available", where this additional knowledge comes into the algorithm just as additional input via the new evaluation function.

Then, for the computation of $a = \sqrt[3]{\pi r^2}$, only three multiplications of the evaluations $r = f(r)$ and $\pi = g(r)$ and one root are required, thus $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \tilde{\Lambda})_{\mathrm{R}} = 0$. However, $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \tilde{\Lambda})_{\mathrm{A}} \geq 1$ still holds since, if there was an Arithmetic finite algorithm for the computation of a from any r, and in particular from $r = 1$, this algorithm yields $\sqrt[3]{\pi}$ after finitely many arithmetic operations. Having this, one easily obtains an Arithmetic algorithm which computes $a/\sqrt[3]{\pi} = \sqrt[3]{\pi r^2}/\sqrt[3]{\pi} = r^{2/3}$ from every input $r > 0$, a contradiction.

Remark 2.11 (Assumptions on Λ). Motivated by the example above, there are several settings that may be considered when analyzing the SCI, for example:

(I) Λ contains all constant functions.
(II) Let Γ be a general algorithm, $A, B \in \Omega$ and $\hat{\Lambda}_\Gamma(A) \subset \Lambda_\Gamma(A)$ denote the set of constant functions. Then $\hat{\Lambda}_\Gamma(A) = \hat{\Lambda}_\Gamma(B)$. In particular, the constant functions are the same for A and B.
(III) Λ contains no constant functions.

Note that, as Example 2.10 demonstrates, given a specific type of tower of algorithms, the SCI may change depending on the assumptions I, II or III. Also, a lower bound on the SCI with assumption I is stronger than a lower bound with assumption II, and similarly with II and III. For upper bounds, the direction is the opposite, an upper bound is stronger with assumption III than with assumption II and similarly with II and I.

When considering general towers of algorithms this does not affect the outcome, as the next theorem establishes. In particular, neither adding constant functions nor canceling them will then change the SCI_G.

Theorem 2.12. *Let $(\Xi, \Omega, \mathcal{M}, \Lambda)$ be a computational problem, and let $\tilde{\Lambda}$ be the union of Λ and all the constant functions on Ω.*

(1) *For every general algorithm $\tilde{\Gamma} : \Omega \to \mathcal{M}$ which, for each $A \in \Omega$, applies certain evaluation functions $\tilde{\Lambda}_{\tilde{\Gamma}}(A) \subset \tilde{\Lambda}$ resp., there exists a general algorithm $\Gamma : \Omega \to \mathcal{M}$ with the same output $\Gamma(A) = \tilde{\Gamma}(A)$ which only applies the evaluation functions from $\Lambda_\Gamma(A) := \tilde{\Lambda}_{\tilde{\Gamma}}(A) \cap \Lambda$, for each $A \in \Omega$, resp.*
(2) $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \tilde{\Lambda})_\mathrm{G} = \mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\mathrm{G}$.

Proof. Obviously, (2) is an immediate consequence of (1): Just replace each general algorithm in the tower by its substitute which is given by (1) and does not make use of any additional constant evaluation functions.

For (1), let $\tilde{\Gamma} : \Omega \to \mathcal{M}$ be a general algorithm. Then $\tilde{\Gamma}(A)$ is determined by $\tilde{\Lambda}_{\tilde{\Gamma}}(A)$ for each $A \in \Omega$, respectively. We just have to show that the output $\tilde{\Gamma}(A)$ can already be identified from $\Lambda_\Gamma(A)$ for each A. Assume the contrary, i.e. there are $A, B \in \Omega$ such that $\tilde{\Gamma}(A) \neq \tilde{\Gamma}(B)$ although $\Lambda_\Gamma(A) = \Lambda_\Gamma(B)$ and $A_f = B_f$ holds for all $f \in \Lambda_\Gamma(A) = \Lambda_\Gamma(B)$. Then we even have $A_f = B_f$ for all $f \in \tilde{\Lambda}_{\tilde{\Gamma}}(A)$ since the functions which additionally occur here are just constants. By condition (iii) in Definition 2.3 $\tilde{\Lambda}_{\tilde{\Gamma}}(B)$ and $\tilde{\Lambda}_{\tilde{\Gamma}}(A)$ must coincide which now easily yields that $\tilde{\Gamma}$ cannot distiguish between A and B, a contradiction. $\square$

We emphasize that by this theorem all the lower bounds on the height of general towers that are proved within this paper are invariant under choosing I, II, or III.

Most of the upper bounds on the SCI are achieved with the more specific arithmetic towers even under assumption III (no constants needed). For the cases where constants are needed we use II, and in those cases the specific constant functions will be spelled out. This is the case for operators with controlled resolvent, bounded dispersion and Schrödinger operators with bounded potential. Interestingly, for Schrödinger operators with potential that blow up at infinity we only need III.

Remark 2.13 (Adaptivity vs. non-adaptivity). One could restrict (i) in Definition 2.3 and require that the evaluation sets $\Lambda_\Gamma(A)$ are the same for all $A \in \Omega$. In this case obviousy (iii) in Definition 2.3 is superfluous. This would give us a *non-adaptive tower of algorithms* as opposed an *adaptive tower of algorithms* where we do not have this restriction. This subtlety is quite important. In particular, it could change the result of the SCI. Note that a bound $\mathrm{SCI} \geq 2$ is stronger if the tower of algorithms is adaptive, however the result $\mathrm{SCI} \leq 2$ gets stronger if the tower of algorithms is required to be non-adaptive. In this article we will always let the towers be adaptive. However, we do want to point out that all our upper bounds for the SCI in this paper are done with non-adaptive towers. Thus, the results that we obtain are actually slightly stronger than what the theorems read.

3. MAIN THEOREMS ON COMPUTING SPECTRA

Computing spectra of linear operators is a fundamental problem that has received an overwhelming amount of attention over the last decades [3–5, 15, 16, 20, 21, 30, 32, 46, 47, 49, 64, 75, 82, 85, 87, 97], and we can only cite a small subset here. We consider the computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ in Example 2.1 (I) in Section 2 with the Hausdorff metric on $\mathcal{M}$ defined by

$$d_{\mathrm{H}}(X,Y) = \max\left\{\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(x,y)\right\}, \tag{3.1}$$

where $d(x,y) = |x-y|$ is the usual Euclidean distance. We ask the basic question:

(i) Given a bounded operator A on a separable Hilbert space $\mathcal{H}$ and suppose that we can access the matrix elements $\{\langle Ae_j, e_i\rangle\}_{i,j\in\mathbb{N}}$, where $\{e_j\}_{j\in\mathbb{N}}$ is some orthonormal basis, can one compute the spectrum of A (given a Radical tower or an Arithmetic tower), and what is the SCI?

This long standing question was partially answered affirmatively in [49], however the result provided was only the bound $\mathrm{SCI}(\Xi,\Omega)_{\mathrm{R}} \leq 3$ where $\Xi(A) = \mathrm{sp}(A)$ and the domain of Ξ is $\Omega = \mathcal{B}(\mathcal{H})$ (we omit the last two variables $\mathcal{M}, \Lambda$ in the SCI when they are obvious). Thus, we were left with a series of open problems such as:

(ii) Is the SCI for spectra of operators > 1 for Arithmetic towers, or even for General towers?
(iii) If so, what is the SCI for spectra of operators in different classes (e.g. self-adjoint, normal, compact, operators with off-diagonal decay etc.)?
(iv) The above questions can be repeated when replacing the spectrum with the pseudo spectrum (see Definition 3.4) or the essential spectrum.

In the following theorems we completely characterize the SCI of spectra and pseudo spectra of different classes of operators and towers.

Remark 3.1. We can think of every separable Hilbert space with an orthonormal basis $\{b_i\}_{i\in\mathbb{N}}$ as the space $l^2(\mathbb{N})$ equipped with the canonical basis $\{e_i\}_{i\in\mathbb{N}}$ in the sense of the isometrical isomorphism given by the mapping $b_i \mapsto e_i$, $i\in\mathbb{N}$. Thus, we can restrict ourselves to this prototype $l^2(\mathbb{N})$ of a separable Hilbert space and the evaluation of all operators w.r.t. the canonical basis in all what follows. In particular, the evaluation set Λ shall be the set of functions $f_{i,j} : A \mapsto \langle Ae_j, e_i\rangle$, $i,j\in\mathbb{N}$, for $A\in\mathcal{B}(l^2(\mathbb{N}))$. We further introduce the orthogonal projections P_n onto the subspaces spanned by $\{e_1,\dots,e_n\}$, respectively.

3.1. Operators with controlled resolvents.

As a start, let $g : [0,\infty) \to [0,\infty)$ be a continuous function, vanishing only at $x=0$ and tending to infinity as $x\to\infty$, and suppose that we know g. We consider bounded operators that have the following property:

$$\|(A-zI)^{-1}\|^{-1} \geq g(\mathrm{dist}(z,\mathrm{sp}(A))) \quad \text{in every point} \quad z\in\mathbb{C}, \tag{3.2}$$

where we use the convention $\|B^{-1}\|^{-1} := 0$ if B has no bounded inverse. Before we prove a result on the SCI of spectra of operators having a common bound of such a type, we commence with some basic remarks. Clearly, we can suppose that $g(x)\leq x$ for all x, since $\mathrm{dist}(z,\mathrm{sp}(A)) \geq \|(A-zI)^{-1}\|^{-1}$ always holds. We further recall that self-adjoint operators and normal operators share this property with $g(x) := x$. Moreover, notice that for every operator A there always exists such a g (define $g(\alpha) := \min\{\|(A-zI)^{-1}\|^{-1} : z\in\mathbb{C} \text{ with } \mathrm{dist}(z,\mathrm{sp}(A)) = \alpha\}$, taking continuity and compactness into account) although there is no g which works for all A. We will address this fact later again.

Remark 3.2 (Assumptions on Λ). In order to make the "additional knowledge" g available for the algorithms we apply the approach of Remark 2.11 and assume that Λ contains, besides the usual evaluations $f_{i,j} : A\mapsto \langle Ae_j, e_i\rangle$ $(i,j\in\mathbb{N})$ also the constant functions $g_{i,j} : A \mapsto g(i/j)$ $(i,j\in\mathbb{N})$, which provide the

values of g in all positive rational numbers. In the case of normal operators we do not need this information, and thus Λ does not need to contain any constant functions.

Theorem 3.3 (Self-adjoint, normal and controlled resolvent). *Given g as above, let Ω_1 denote the set of all bounded operators on $l^2(\mathbb{N})$ with the property* (3.2), *Ω_2 the set of all bounded normal operators and Ω_3 the set of all bounded self-adjoint operators. Consider*

$$\Xi : \Omega_i \ni A \mapsto \mathrm{sp}(A) \in \mathcal{M} \quad i = 1, 2, 3.$$

Then, for each $i = 1, 2, 3$,

$$\mathrm{SCI}(\Xi, \Omega_i)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega_i)_{\mathrm{A}} = 2. \tag{3.3}$$

Moreover, (3.3) *holds if Ω_i is replaced by Ω_i^M where $\Omega_i^M := \{A \in \Omega_i : \|A\| \leq M\}$, where $M > 0$.*

The fact that the SCI of spectra of self-adjoint operators is equal to two may come as a surprise. In particular, even when faced with a problem that depends continuously on the input, there is no tower that can compute the spectrum in one limit. In fact, since $\mathrm{SCI}(\Xi, \Omega_i)_{\mathrm{G}} = 2$, adding more operations allowed in the tower will never solve this issue. In particular, even if one could compute the spectrum of a finite-dimensional matrix using finitely many arithmetic operations (of course no such algorithm exists), it is still impossible to compute spectra of self-adjoint operators in one limit. However, the good news that $\mathrm{SCI}(\Xi, \Omega_i)_{\mathrm{A}} = 2$, thus, the computation can be done in two limits when arithmetic operations are allowed. Note how this improves the results from [49] substantially.

3.2. **Non-normal operators.** Another set that has received substantial attention is the pseudo spectrum. This set has been popular in spectral theory, analysis of pseudo differential operators and non-Hermitian quantum mechanics. Recently, the concept of pseudo spectra has been generalized to what one refers to as N-pseudo spectra.

Definition 3.4 (Pseudospectra). *For $N \in \mathbb{Z}_+$ and $\epsilon > 0$, the (N, ϵ)-pseudospectrum of a bounded linear operator $A \in \mathcal{B}(\mathcal{H})$ on a Hilbert space $\mathcal{H}$ is defined as the set*

$$\mathrm{sp}_{\mathrm{N},\epsilon}(A) := \{z \in \mathbb{C} : \|(A - zI)^{-2^N}\|^{2^{-N}} \geq 1/\epsilon\}.$$

For $N = 0$ this is the (classical) ϵ-pseudospectrum

$$\mathrm{sp}_{\epsilon}(A) := \{z \in \mathbb{C} : \|(A - zI)^{-1}\| \geq 1/\epsilon\}.$$

For more information on pseudospectra we refer to [33, 46, 47, 49, 82, 86, 97]. Also recall that the sets $\mathrm{sp}_{\mathrm{N},\epsilon}(A)$ are continuous w.r.t. the parameter $\epsilon > 0$, and converge to $\mathrm{sp}(A)$ as $\epsilon \to 0$ for every A. Another class of operators of interest is the family of operators with controllable off-diagonal decay. Before we get into the details we must introduce the concept of dispersion.

Definition 3.5 (Dispersion). *We say that the dispersion of $A \in \mathcal{B}(l^2(\mathbb{N}))$ is bounded by the function $f : \mathbb{N} \to \mathbb{N}$ if*

$$D_{f,m}(A) := \max\{\|(I - P_{f(m)})AP_m\|, \|P_m A(I - P_{f(m)})\|\} \to 0 \quad \textit{as } m \to \infty.$$

Note that for every operator A there is always a function f which is a bound for its dispersion since AP_m, $P_m A$ are compact and (P_n) converges strongly to the identity. But there is no function f which acts as a uniform bound for all operators. Nevertheless, there are important (sub)classes of operators having well known uniform bounds, which should be mentioned:

(i) band operators with bandwidth less than d: $f(k) = k + d$.

(ii) band-dominated and weakly band-dominated operators: $f(k) = 2k$. For definitions and properties of band and band-dominated operators see [63, 77, 83]. Weakly band-dominated operators can be found in [66].

(iii) Laurent/Toeplitz operators with piecewise continuous generating function: $f(k) = k^2$ (cf. [17] and [56, Proposition 5.4]).

(iv) Let $\mathcal{F}$ be a family of bounded operators with a common bound f. Then $\tilde{f}$, given by $\tilde{f}(k) = f(k)+k$, is a common bound for all operators in the Banach algebra which is generated by $\mathcal{F}$.

Remark 3.6 (Assumptions on Λ). In the case when the dispersion of the operator is known, the values $f(m)$ $(m \in \mathbb{N})$ shall be available to the algorithms as constant evaluation functions (as was done in Remark 2.11 and also for the function g above). However, if the dispersion is not known then Λ will not contain any constant functions in the theorems below.

Theorem 3.7 (General, compact and dispersion operators). *Define the following primary sets: $\Omega_1 = \mathcal{B}(l^2(\mathbb{N}))$, for $f : \mathbb{N} \to \mathbb{N}$ let Ω_2 denote the set of bounded operators on $l^2(\mathbb{N})$ whose dispersion is bounded by f and $\Omega_3 = \mathcal{K}(l^2(\mathbb{N}))$, where $\mathcal{K}(l^2(\mathbb{N}))$ denotes the set of all compact operators. Define the following problem functions: $\Xi_1(A) = \mathrm{sp}(A)$ and for $\epsilon > 0$ and $N \in \mathbb{Z}_+$, let $\Xi_2(A) = \mathrm{sp}_{\mathrm{N},\epsilon}(A)$. Then*

$$\mathrm{SCI}(\Xi_1,\Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1,\Omega_1)_{\mathrm{A}} = 3, \qquad \mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{A}} = 2, \tag{3.4}$$

$$\mathrm{SCI}(\Xi_1,\Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1,\Omega_2)_{\mathrm{A}} = 2, \qquad \mathrm{SCI}(\Xi_2,\Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2,\Omega_2)_{\mathrm{A}} = 1, \tag{3.5}$$

$$\mathrm{SCI}(\Xi_1,\Omega_3)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1,\Omega_3)_{\mathrm{A}} = 1, \qquad \mathrm{SCI}(\Xi_2,\Omega_3)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2,\Omega_3)_{\mathrm{A}} = 1. \tag{3.6}$$

Moreover, (3.4), (3.5) *and* (3.6) *hold if Ω_i is replaced by Ω_i^M where $\Omega_i^M = \{A \in \Omega_i : \|A\| \leq M\}$, where $M > 0$.*

By Theorems 3.3 and 3.7 we observe that the SCI is equal to three as long as there is no additional information on the structure of the operators under consideration known and can be taken into account. Once a function g is available which estimates the behaviour of the resolvent norm in the sense of (3.2), the SCI decreases to two . The same holds true, if a bound f on the dispersion is known.

3.3. **Computing the essential spectrum.** We will end this section with the intriguing result that in the case of operators with g-bounded resolvent norm and f-bounded dispersion, we have that the SCI for the spectrum is indeed one, yet the SCI for the essential spectrum is two. This may come as a surprise, however, what this result tells us is that the problem of distinguishing between the essential spectrum and the isolated eigenvalues with finite multiplicity cannot be solved by an algorithm using only one limit. In other words: one can compute the spectrum in one limit, but one will never be able to distinguish the essential spectrum from the total spectrum. This is summed up in the following theorem.

Theorem 3.8 (Spectrum vs. essential spectrum). *For $A \in \Omega_1 = \mathcal{B}(l^2(\mathbb{N}))$ define $\Xi_1(A) = \mathrm{sp}(A)$ and $\Xi_2(A) = \mathrm{sp}_{\mathrm{ess}}(A)$ (the essential spectrum). Let $f : \mathbb{N} \to \mathbb{N}$, $g : [0,\infty) \to [0,\infty)$ and let Ω_2 denote the set of bounded operators with the property* (3.2) *(or self-adjoint or normal operators) on $l^2(\mathbb{N})$ whose dispersion is bounded by f. Then*

$$\begin{aligned} &\mathrm{SCI}(\Xi_1,\Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1,\Omega_2)_{\mathrm{A}} = 1, \\ &\mathrm{SCI}(\Xi_2,\Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2,\Omega_2)_{\mathrm{A}} = 2, \\ &\mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{A}} = 3. \end{aligned} \tag{3.7}$$

Moreover, all the results in (3.7) hold if Ω_i is replaced by Ω_i^M where $\Omega_i^M = \{A \in \Omega_i : \|A\| \leq M\}$, where $M > 0$.

Having considered bounded operators in this section we now turn to arguably one of the most important unbounded operators in the last decades: the Schrödinger operator.

4. Main theorems on computational quantum mechanics

The Schrödinger operator

$$H = -\Delta + V, \qquad V : \mathbb{R}^d \to \mathbb{C}, \tag{4.1}$$

is one of the most celebrated operators of modern times and the mainstay in quantum mechanics, and it is its spectrum that is of utmost importance. If we fix the domain of H such that it is appropriate for a class of potentials V, the spectrum of H is uniquely determined by V. The basic question is therefore:

> Given the potential V, so that V can be evaluated at any point $x \in \mathbb{R}^d$, can one compute the spectrum (or pseudospectrum) of H, and what is the SCI?

Note that this problem has been unsolved for a long time when considering H acting on $\mathrm{L}^2(\mathbb{R}^d)$ allowing non self-adjointness and arbitrary complex potentials. This is not a surprise given that the problem of computing spectra of arbitrary non self-adjoint infinite matrices has only recently been solved [49]. However, there is a vast amount of work on how to compute spectra of Schrödinger operators with specific real potentials [10, 18, 19, 27, 29, 36, 65, 73]. We emphasise the importance of generality as we want the theory to include non-Hermitian quantum mechanics [8,9,51,52] and the theory of resonances [92,101]. The key issue is that the only available input is the potential function V (no matrix values are assumed). In this section we shall solve this problem for very general classes of potentials and give bounds on the SCI. We consider the computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ in Example 2.1 (IV) in Section 2 with the *Attouch-Wets* metric defined by

$$d_{\mathrm{AW}}(A, B) = \sum_{i=1}^{\infty} 2^{-i} \min\left\{1, \sup_{|x|<i} |d(x, A) - d(x, B)|\right\}, \tag{4.2}$$

where A and B are closed subsets of $\mathbb{C}$, and where $d(x, A)$ is the usual Euclidean distance between the point $x \in \mathbb{C}$ and A, which is well-defined even when A is unbounded. We discuss some properties of this metric in Remark 11.2. Also, since the pseudospectrum may be discontinuous w.r.t ϵ in the case of unbounded operators, it is more convenient to redefine it for the unbounded operator H to be

$$\mathrm{sp}_\epsilon(H) := \mathrm{cl}(\{z \in \mathbb{C} : \|(H - zI)^{-1}\| > 1/\epsilon\}).$$

This is however equivalent to the definition in the bounded case in Definition 3.4 (see e.g. [14, 42, 86]).

4.1. Bounded potentials.

We will first consider

$$\Omega_1 := \{V : V \in \mathrm{L}^\infty(\mathbb{R}^d) \cap \mathrm{BV}_\phi(\mathbb{R}^d)\},$$

where $\phi : [0, \infty) \to [0, \infty)$ is some increasing function and

$$\mathrm{BV}_\phi(\mathbb{R}^d) = \{f : \mathrm{TV}(f_{[-a,a]^d}) \le \phi(a)\}, \tag{4.3}$$

($f_{[-a,a]^d}$ means f restricted to the box $[-a, a]^d$) with TV being the total variation of a function in the sense of Hardy and Krause (see [70]).

Also, we consider Schrödinger operators with controlled resolvents. In particular, let $g : [0, \infty) \to [0, \infty)$ be a continuous function, vanishing only at $x = 0$ and tending to infinity as $x \to \infty$, and suppose that we know g. Define

$$\Omega_2 = \Omega_1 \cap \tilde{\Omega}, \quad \tilde{\Omega} = \{V : \|(-\Delta + V - zI)^{-1}\|^{-1} \ge g(\mathrm{dist}(z, \mathrm{sp}(-\Delta + V)))\}. \tag{4.4}$$

Note that the set Ω_1 requires a little bit more than V just being locally of bounded variation. We also need to know an upper bound on the growth of the total variation as we restrict the function to a larger set. In

particular we need to know ϕ or an estimate for it. The set $\tilde{\Omega}$ of operators obviously includes self-adjoint and normal operators, however, is much bigger.

Remark 4.1 (Assumptions on Λ). As done in the case of bounded Hilbert space operators, and as discussed in Remark 2.11, the additional knowledge of g is available for the algorithms by assuming that Λ also contains the constant functions $g_{i,j} : V \mapsto g(i/j)$ $(i, j \in \mathbb{N})$, which provide the values of g in all positive rational numbers. In addition we will assume that Λ contains certain constant functions that will be specified in Section 11.1 in Remark 11.6 .

Theorem 4.2 (Bounded potential). *Let $\mathcal{D}(H) = \mathrm{W}^{2,2}(\mathbb{R}^d)$ and define $\Xi_1(V) = \mathrm{sp}(-\Delta + V)$ and, for $\epsilon > 0$, let $\Xi_2(V) = \mathrm{sp}_\epsilon(-\Delta + V)$. Then*

$$\mathrm{SCI}(\Xi_1, \Omega)_{\mathrm{A}} \begin{cases} \leq 2 & \Omega = \Omega_1 \\ = 1 & \Omega = \Omega_2 \end{cases}, \qquad \mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{A}} \leq 2.$$

Remark 4.3. As will be evident from the proof techniques, one can build towers of algorithms for operators with more general classes of potentials (for example $\mathrm{L}^\infty(\mathbb{R}^d) \cap \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)$) or $\mathrm{L}^2(\mathbb{R}^d) \cap \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)$), however, the height of these towers will be higher than the ones considered in this paper. The main future task is to obtain exact values of the SCI of the spectrum given the different potential classes.

4.2. **Unbounded potentials.** We get a rather intriguing phenomenon for sectorial operators. Namely, the SCI of both the spectrum and the pseudospectrum is one. In particular, suppose that we have nonnegative θ_1, θ_2 such that $\theta_1 + \theta_2 < \pi$. Define

$$\Omega = \{V \in \mathrm{C}(\mathbb{R}^d) : \forall x \arg(V(x)) \in [-\theta_2, \theta_1], |V(x)| \to \infty \text{ as } x \to \infty\}. \tag{4.5}$$

We define the operator H via the minimal operator h as: $H = h^{**}$, $h = -\Delta + V$, $\mathcal{D}(h) = \mathrm{C}_c^\infty(\mathbb{R}^d)$. When $V \in \Omega$ it follows that H has compact resolvent, a result that we also establish as a part of the proof of the following theorem.

Remark 4.4 (Assumptions on Λ). Interestingly, no constant function are needed in Λ in order to obtain the results in the following theorem, as opposed to the case where we have a bounded potential.

Theorem 4.5 (Unbounded potential). *Let $\Xi_1 : \Omega \ni V \mapsto \mathrm{sp}(H)$ and, for $\epsilon > 0$, $\Xi_2 : \Omega \ni V \mapsto \mathrm{sp}_\epsilon(H)$. Then*

$$\mathrm{SCI}(\Xi_1, \Omega)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1, \Omega)_{\mathrm{A}} = \mathrm{SCI}(\Xi_2, \Omega)_{\mathrm{G}} = \mathrm{SCI}(\Xi_2, \Omega)_{\mathrm{A}} = 1.$$

Note that the key to this result is the compact resolvent of H. The fact that the SCI is one in this case is most natural as the SCI for spectra and pseudo spectra of compact operators is one. The continuity assumption on V in Theorem 4.5 is to make sure that the discretization used actually converges. However, by tweaking with the approximation this assumption can may be weakened to include potentials that have certain discontinuities.

5. Main theorems on solving linear systems

Just as finding spectra of operators and roots of polynomials, the problem of solving linear systems of equations is at the heart of computational mathematics. For the finite-dimensional case it is easy to find an algorithm that can perform the task, but what about the infinite-dimensional case? In particular, if $b \in l^2(\mathbb{N})$, $A \in \mathcal{B}_{\mathrm{inv}}(l^2(\mathbb{N}))$ (the set of bounded invertible operators) and $\Omega \subset \mathcal{B}_{\mathrm{inv}}(l^2(\mathbb{N})) \times l^2(\mathbb{N})$ and we define the mapping $\Xi : \Omega \ni (A, b) \mapsto A^{-1}b$, what is the SCI of this mapping given different domains Ω?

Remark 5.1 (Assumptions on Λ). Here, as in Example 2.1, we again suppose that the set Λ of evaluations consists of the functions which read the matrix elements $\{\langle Ae_j, e_i\rangle\}_{i,j\in\mathbb{N}}$ and the sequence entries $\{\langle b, e_k\rangle\}_{k\in\mathbb{N}}$ of $(A, b) \in \Omega$. Also, in the case when the dispersion of the operator is known, the values $f(m)$ $(m \in \mathbb{N})$ shall be available to the algorithms as constant evaluation functions (as was done in Remark 2.11). However, if the dispersion is not known then Λ will not contain any constant functions in the theorems below.

Theorem 5.2 (Linear systems). *Let $\mathcal{B}_{\mathrm{inv},f}(l^2(\mathbb{N}))$ denote the set of bounded invertible operators with dispersion bounded by $f : \mathbb{N} \to \mathbb{N}$, $\mathcal{B}_{\mathrm{inv},sa}(l^2(\mathbb{N}))$ denote the set of bounded invertible self-adjoint operators, and define the domains $\Omega_1 = \mathcal{B}_{\mathrm{inv}}(l^2(\mathbb{N})) \times l^2(\mathbb{N})$, $\Omega_2 = \mathcal{B}_{\mathrm{inv},sa}(l^2(\mathbb{N})) \times l^2(\mathbb{N})$ and $\Omega_3 = \mathcal{B}_{\mathrm{inv},f}(l^2(\mathbb{N})) \times l^2(\mathbb{N})$. Then*

$$\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} = 2, \tag{5.1}$$

$$\mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{A}} = 2, \tag{5.2}$$

$$\mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{A}} = 1. \tag{5.3}$$

Another problem of interest when dealing with solutions of linear systems of equations is the computation of the norm of the inverse. This is obviously related to the stability of the problem. The task of computing the norm of the inverse of an operator can also be analysed in terms of the SCI, and that is the topic of the next theorem.

Theorem 5.3 (Computing norm of the inverse). *Let $\Omega_1 = \mathcal{B}(l^2(\mathbb{N}))$, Ω_2 the subset of self-adjoint operators, Ω_3 the subset of operators with dispersion bounded by an $f : \mathbb{N} \to \mathbb{N}$, and let $\Xi : A \mapsto \|A^{-1}\|$.* [1] *Then*

$$\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} = 2, \tag{5.4}$$

$$\mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{A}} = 2, \tag{5.5}$$

$$\mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{A}} = 1. \tag{5.6}$$

6. Main theorems on the impossibility of error control

One of the natural desires in computations is error control. In particular, given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ with $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_\alpha$ = k for some tower of algorithms of type α, and a tower of algorithms of height k, $\Gamma_{n_k}, \dots, \Gamma_{n_k,\dots,n_1}$, it is highly desirable to be able to control the convergence $\Gamma_{n_k} \to \Xi, \dots, \Gamma_{n_k,\dots,n_1} \to \Gamma_{n_k,\dots,n_2}$. More precisely, for $\epsilon > 0$, how big do $n_k, \dots, n_1$ have to be such that

$$d(\Gamma_{n_k,\dots,n_1}(A), \Xi(A)) \le \epsilon, \qquad \forall\, A \in \Omega.$$

This type of global error control is quite common in different areas of computational theory such as in differential equations and integration [55].

Unfortunately, such choices of $n_k, \dots, n_1$ may be impossible. More precisely, problems with SCI greater than one with respect to a General tower will never have error control. This is summed up in the following theorem.

Theorem 6.1 (No global error control). *Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ where we have $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} \ge 2$. Suppose that there is a general tower of algorithms of height k, $\Gamma_{n_k}, \dots, \Gamma_{n_k,\dots,n_1}$ for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$. Then there do NOT exist integers $n_k = n_k(m), \dots, n_1 = n_1(m)$ (depending on m) such that*

$$d(\Gamma_{n_k,\dots,n_1}(A), \Xi(A)) \le \frac{1}{m}, \qquad \forall\, A \in \Omega, \quad \forall\, m \in \mathbb{N}.$$

[1] As usual, $\|A^{-1}\| := \infty$ if A is not invertible.

With such a negative result, it is natural to consider a weaker concept than global error control such as local error control. In particular one could ask if the following is true:

$$\forall A \in \Omega \text{ and } \forall \epsilon > 0,\ \exists n_k, \dots, n_1 \text{ such that } d(\Gamma_{n_k,\dots,n_1}(A), \Xi(A)) < \epsilon.$$

Indeed, it is, and note that the existence of $n_k, \dots, n_1$ is guaranteed by the definition of a Tower of algorithms. If we could find the integers $n_k, \dots, n_1$ we would call this *local error control*. However, the integers $n_k, \dots, n_1$ cannot be computed as the next theorem demonstrates, and thus local error control is also impossible.

Theorem 6.2 (Local error control cannot be computed). *Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ with $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} \geq 2$, suppose that there is a general tower of algorithms $\Gamma_{n_k}, \dots, \Gamma_{n_k,\dots,n_1}$ of height k for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$. Then, there does NOT exist a sequence $\{\tilde{\Gamma}_n\}$ of general algorithms $\tilde{\Gamma}_n : \Omega \to \mathbb{N}^k$ such that for any $A \in \Omega$,*

$$d(\Gamma_{\tilde{\Gamma}_n(A)_k,\dots,\tilde{\Gamma}_n(A)_1}(A), \Xi(A)) < \frac{1}{n}.$$

Remark 6.3. The conclusion of Theorem 6.1 and Theorem 6.2 is the following. Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ with $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} \geq 2$, then no one can ever know when one is "epsilon" away from the solution of the problem. In particular, if we have that $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} = \mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{A}} = k \geq 2$, one can compute an approximation to the solution, but never know when to stop.

A weaker requirement than error control would be that one could reindex the tower to get only one limit. In particular, let $\Gamma_{n_k}, \dots, \Gamma_{n_k,\dots,n_1}$ be a General tower of algorithms for a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ where the cardinality of Λ is infinite. By (2.2), for any $A \in \Omega$, there exists a reindexing function $h_A : \mathbb{N} \to \mathbb{N}^k$ such that

$$\Gamma_{h_A(n)_k,\dots,h_A(n)_1} \longrightarrow \Xi(A), \qquad n \to \infty.$$

However, the function h_A is also impossible to compute as the following theorem establishes.

Theorem 6.4 (Impossibility of computing the reindexing function). *Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ with $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} \geq 2$, suppose that there is a general tower $\Gamma_{n_k}, \dots, \Gamma_{n_k,\dots,n_1}$ of algorithms of height k for $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$. Then, there does NOT exist a sequence $\{\tilde{\Gamma}_n\}$ of general algorithms $\tilde{\Gamma}_n : \Omega \to \mathbb{N}^k$ such that for any $A \in \Omega$,*

$$\Gamma_{\tilde{\Gamma}_n(A)_k,\dots,\tilde{\Gamma}_n(A)_1}(A) \longrightarrow \Xi(A), \qquad n \to \infty.$$

Remark 6.5. Note that the result above sparks the discussion about which problems can actually have error control.

(i) **(Polynomial root finding with rational maps)** As we will see in Section 8, the SCI of finding a root of the quintic given a Doyle-McMullen tower is greater than one. However, this does not imply that one cannot have error control. In particular, a potential reindexing of the functions in the tower, due to error control, could be possible. The new reindexed tower, however, will not be a Doyle-McMullen tower. Thus, the reindexing does not violate the bound of the SCI w.r.t. Doyle-McMullen towers. The situation is very different in the case of inverse problems.

(ii) **(Inverse Problems)** As we discussed in Section 5 the SCI for solving arbitrary linear systems with a General tower is two. This is a much stronger statement when it comes to limitations of error control than in the polynomial case. In particular, since this holds for an arbitrary General tower, one cannot provide error control regardless of what kind of mathematical tools are allowed. However, the interesting problem will be to determine which subclasses of problems can be solved with error

control. These problems must have $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} \leq 1$, however, this is of course only a necessary but not sufficient criterion.

7. Computability, Arithmetical Hierarchy, mathematical logic and the SCI

In this section we build a bridge to the classical theory of computability and mathematical logic, and demonstrate how the SCI and towers of algorithms extend important concepts in that field. Decision making problems are at the heart of the theory of computation and as a motivation we start with some basic problems. Intriguingly, it is through this framework we can prove some of the lower bounds on the SCI for spectral problems.

7.1. **Decision making.** Within this section we exclusively deal with problems (functions)

$$\Xi : \Omega \to \mathcal{M} := \{\textit{Yes}, \textit{No}\},$$

where $\mathcal{M}$ is equipped with the discrete metric. This means that for such problems we search for General algorithms $\Gamma_{n_k,\ldots,n_1} : \Omega \to \mathcal{M}$ which, for a given input $\omega \in \Omega$, answer *Yes* or *No*. We will refer to such problems as decision making problems. Clearly, a sequence $\{m_i\} \subset \mathcal{M}$ of such "answers" converges to $m \in \mathcal{M}$ if and only if finitely many m_i are different from m. The number of decision problems is countless, and we will only consider some basic illustrative problems here.

Let Ω_1 denote the collection of all sequences $\{a_i\}_{i\in\mathbb{N}}$ with entries $a_i \in \{0, 1\}$. For $\{a_i\}_{i\in\mathbb{N}} \in \Omega_1$ we define $\Xi(\{a_i\})$ to be the answer to the following question,

$\Xi_1(\{a_i\})$: Does $\{a_i\}_{i\in\mathbb{N}}$ have a non-zero entry?

For such problems the evaluation set Λ shall consist of the functions $f_k : \{a_i\} \mapsto a_k$, $k \in \mathbb{N}$, which read the kth entry of a given function $\{a_i\}_{i\in\mathbb{N}}$, respectively. It is easy to see that $\mathrm{SCI}(\Xi_1, \Omega_1)_{\mathrm{G}} = 1$ in this case, however, we could ask the slightly more difficult question

$\Xi_2(\{a_i\})$: Does $\{a_i\}$ have infinitely many non-zero entries?

And of course, we could extend these questions to matrices. In particular, let Ω_2 denote the collection of all infinite matrices $\{a_{i,j}\}_{i,j\in\mathbb{N}}$ with entries $a_{i,j} \in \{0, 1\}$, and consider the following problems,

$\Xi_3(\{a_{i,j}\})$: Does $\{a_{i,j}\}$ have a non-zero entry?
$\Xi_4(\{a_{i,j}\})$: Does $\{a_{i,j}\}$ have infinitely many non-zero entries?

The situation becomes more complex and less clear with the following questions:

$\Xi_5(\{a_{i,j}\})$: Does $\{a_{i,j}\}$ have a column containing infinitely many non-zero entries?
$\Xi_6(\{a_{i,j}\})$: Does $\{a_{i,j}\}$ have infinitely many columns containing infinitely many non-zero entries?
$\Xi_7(\{a_{i,j}\})$: Does $\{a_{i,j}\}$ have (only) finitely many columns with (only) finitely many 1s?

And finally the following question that intriguingly is linked to the problem of showing the lower bound on SCI for spectra of operators. In this case we consider infinite matrices indexed by $\mathbb{Z}$ rather than $\mathbb{N}$. In particular, let Ω_3 denote the collection of all infinite matrices $\{a_{i,j}\}_{i,j\in\mathbb{Z}}$ with entries $a_{i,j} \in \{0, 1\}$. We then ask the following question.

$$\Xi_8 : \{a_{i,j}\}_{i,j\in\mathbb{Z}} \mapsto \left(\exists D \forall j \left(\left(\forall i \sum_{k=-i}^{i} a_{k,j} < D\right) \vee \left(\forall R \exists i \sum_{k=0}^{i} a_{k,j} > R \wedge \sum_{k=-i}^{0} a_{k,j} > R\right)\right)\right)$$

("there is a bound D such that every column has either less than D 1s or is two-sided infinite")

The purpose of this section is to first demonstrate how decision problems, that are the core of classical recursion theory, fit naturally into the SCI framework. And also provide key results that are crucial for proving lower bounds on the SCI for other problems. The following theorem sums up the bounds for the SCI for the first four of these problems.

Theorem 7.1 (Decision making problems 1-4). *Given the setup above we have*

$$\begin{aligned}\mathrm{SCI}(\Xi_1,\Omega_1)_{\mathrm{G}} &= \mathrm{SCI}(\Xi_1,\Omega_1)_{\mathrm{A}} = 1,\\ \mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{G}} &= \mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{A}} = 2,\\ \mathrm{SCI}(\Xi_3,\Omega_2)_{\mathrm{G}} &= \mathrm{SCI}(\Xi_3,\Omega_2)_{\mathrm{A}} = 1,\\ \mathrm{SCI}(\Xi_4,\Omega_2)_{\mathrm{G}} &= \mathrm{SCI}(\Xi_4,\Omega_2)_{\mathrm{A}} = 2,\end{aligned}$$

Theorem 7.2 (Decision making problems 5-8). *Given the setup above we have*

$$\begin{aligned}\mathrm{SCI}(\Xi_5,\Omega_2)_{\mathrm{G}} &= \mathrm{SCI}(\Xi_5,\Omega_2)_{\mathrm{A}} = 3,\\ \mathrm{SCI}(\Xi_7,\Omega_2)_{\mathrm{G}} &= \mathrm{SCI}(\Xi_7,\Omega_2)_{\mathrm{A}} = 3,\\ \mathrm{SCI}(\Xi_8,\Omega_3)_{\mathrm{G}} &= \mathrm{SCI}(\Xi_8,\Omega_3)_{\mathrm{A}} = 3,\\ 3 \leq \mathrm{SCI}(\Xi_6,\Omega_2)_{\mathrm{G}} &\leq \mathrm{SCI}(\Xi_6,\Omega_2)_{\mathrm{A}} \leq 4.\end{aligned}$$

We may also look at spectral theory from a decision making point of view and ask: How difficult is it to decide whether a given point $\lambda \in \mathbb{C}$ belongs to the spectrum of a given bounded linear operator A or not? Obviously this question is equivalent to asking whether $0 \in \mathrm{sp}(A - \lambda I)$. Thus, it suffices to study the question for $\lambda = 0$. To make this more precise, we consider the functions

$$\mathrm{sp}^0 : \mathcal{B}(l^2(\mathbb{N})) \to \mathcal{M} := \{\textit{Yes}, \textit{No}\}, \quad A \mapsto (0 \in \mathrm{sp}(A))$$

The following theorem addresses this question.

Theorem 7.3 (Decision making and spectra). *Let $\Omega \subset \mathcal{B}(l^2(\mathbb{N}))$ and define $\Xi : \Omega \ni A \mapsto \mathrm{sp}^0(A)$. Let $\Omega_1 = \mathcal{B}(l^2(\mathbb{N}))$, $f : \mathbb{N} \to \mathbb{N}$ and Ω_2 denote the set of bounded operators on $l^2(\mathbb{N})$ whose dispersion is bounded by f, and finally let Ω_3 denote the set of bounded self-adjoint diagonal operators. Then*

$$\begin{aligned}\mathrm{SCI}(\Xi,\Omega_1)_{\mathrm{G}} &= \mathrm{SCI}(\Xi,\Omega_1)_{\mathrm{A}} = 3,\\ \mathrm{SCI}(\Xi,\Omega_2)_{\mathrm{G}} &= \mathrm{SCI}(\Xi,\Omega_2)_{\mathrm{A}} = 2,\\ \mathrm{SCI}(\Xi,\Omega_3)_{\mathrm{G}} &= \mathrm{SCI}(\Xi,\Omega_3)_{\mathrm{A}} = 2.\end{aligned}$$

This result reveals that even for self-adjoint diagonal operators there do not exist any height-one towers of algorithms that can decide sp^0, although there is a height-one tower which computes the whole spectrum as is claimed in Theorem 3.8. This seems to be a bit surprising at a first glance. Actually, the present question is really stronger in a sense: From Theorem 3.8 we only get approximations for the spectrum which converge with respect to the Hausdorff distance, but which can still have even an empty intersection with $\mathrm{sp}(A)$, whereas Theorem 7.3 addresses the inclusion $\lambda \in \mathrm{sp}(A)$.

Note that the SCI of the decision problems above are considered with respect to general and arithmetic towers. These towers do not assume any computability model, but only a model on the mathematical tools allowed (arithmetic operations in the case of arithmetic tower) and the way the algorithm can read information (only finite amount of input). However, the SCI framework with towers of algorithms fit naturally into the classical theory of computability and the Arithmetical Hierarchy.

7.2. **Computable functions.** Our considerations will be based on so-called *recursive functions* whose definition is due to Gödel and can be found e.g. in [72, I.1.7]. The precise definition is not really important for our following considerations since, due to a *Basic Theorem* ([72, Theorem I.7.12]), it is equivalent to Turing computability, register machine computability, being finitely definable, Church's λ-definability, flowchart computability, or many other well-established concepts of computable functions (see also [60, 61] and a recent survey article on the topic [34]). *Church's Thesis* (Church, Turing, 1936) proposes that

Every effectively computable function is recursive.

Therefore, in all what follows we will refer to recursive functions simply as *computable functions* meaning one or all of the above precise concepts, or by Church's Thesis any 'effectively computable function'. For a different model based on the Blum-Shub-Smale machine see [22]

Remark 7.4. Functions $R : \mathbb{Z}_+^n \to \{\mathit{true}, \mathit{false}\}$ (or $\{\mathit{Yes}, \mathit{No}\}$ in our previous notation) are called n-ary relations. Note that such relations can also be regarded as the characteristic functions of the respective sets $\{(x_1, \ldots, x_n) \in \mathbb{Z}_+^n : R(x_1, \ldots, x_n)\}$.

7.3. **The Arithmetical Hierarchy.** A corner stone in classical logic is the notion of Arithmetical Hierarchy, which we will see later is strongly connected to the SCI and towers of algorithms. Before we can present our main contribution in this section, we need to recall some basic definitions and results.

Definition 7.5 (Arithmetical Hierarchy). *A relation R is in the Arithmetical Hierarchy if R is computable or if there exists a computable relation S such that R can be obtained from S by some finite sequence of complementation and/or projection operations. For this, the projection of $f : \mathbb{Z}_+^n \to \{\mathit{true}, \mathit{false}\}$ along the jth coordinate is defined as the characteristic function of the set (which is a subset of $\mathbb{Z}_+^{n-1}$)*

$$\{(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) : (\exists x_j) f(x_1, \ldots, x_n)\}.$$

The following equivalent conditions describe the Arithmetical Hierarchy (see [72, IV.1] and [79, 14.1-4]). In particular, let R be an n-ary relation, then the following are equivalent:

(i) R is in the Arithmetical Hierarchy.

(ii) R can be represented as

$$(x_1, \ldots, x_n) \mapsto (Q_1 y_1) \cdots (Q_m y_m) S(x_1, \ldots, x_n, y_1, \ldots, y_m)$$

where Q_i is either $(\forall)$ or $(\exists)$ for $i = 1, \ldots, m$, and S is an $(n+m)$-ary computable relation.

(iii) R can be represented as

$$(x_1, \ldots, x_n) \mapsto (Q_1 y_1) \cdots (Q_k y_k) T(x_1, \ldots, x_n, y_1, \ldots, y_k)$$

where (Q_i) is a list of alternated quantifiers ($(\forall)$ and $(\exists)$), and T is an $(n+k)$-ary computable relation (Prenex normal form, Kuratowsky, Tarski 1931).

(iv) R is definable in First-Order Arithmetic (Gödel, 1936).

Following [62] we define the classes of Σ_n, Π_n and Δ_n relations, proceeding by induction:

Definition 7.6 (Σ_m, Π_m, Δ_m). *Let $m \in \mathbb{Z}_+$. We then define the following.*

(i) *A relation is Σ_0 and Π_0 if it is computable.*

(ii) *A relation is Σ_{m+1} if it can be expressed in the form $(\exists y) S(x, y)$, where $S(x, y)$ is Π_m.*

(iii) *A relation R is Π_{m+1} if its complementary relation $\neg R$ is Σ_{m+1}.*

(iv) $\Delta_m := \Sigma_m \cap \Pi_m$.

It is easily seen that a relation $R(x)$ is Σ_m iff it has a definition of the form

$$(x_1, \ldots, x_n) \mapsto (\exists y_1)(\forall y_2) \cdots S(x_1, \ldots, x_n, y_1, \ldots, y_m),$$

where $S(x, y)$ is computable and there are m alternating quantifiers starting with $\exists$. An analogous observation holds for Π_m relations, with m alternating quantifiers starting with $\forall$. This hierarchy is the *Arithmetical Hierarchy*, or *Kleene-Mostowski Hierarchy* [60, 69], and does not collapse. More precisely, we have the following [72, IV.1.13]:

Theorem 7.7 (Hierarchy theorem). *For any $m \in \mathbb{N}$, we have the following:*

(i) $\Sigma_m \setminus \Pi_m \neq \emptyset$, *hence* $\Delta_m \subsetneq \Sigma_m$.

(ii) $\Pi_m \setminus \Sigma_m \neq \emptyset$, *hence* $\Delta_m \subsetneq \Pi_m$.
(iii) $\Sigma_m \cup \Pi_m \subsetneq \Delta_{m+1}$.
(iv) $R \in \bigcup_{n\in\mathbb{Z}_+} \Sigma_n$ *if and only if* R *is in the Arithmetical Hierarchy.*

Note that this hierarchy is related to the classification based on the Turing jump operation by Post's Theorem (e.g. [79, 14.5 Theorem VIII]). See also [100] regarding the Arithmetical Hierarchy of real numbers.

7.4. **The SCI and the Arithmetical Hierarchy.** Given a subset $A \subset \mathbb{Z}_+$ with characteristic function χ_A being definable in First-Order Arithmetic, we are interested in the SCI of deciding whether a given number $x \in \mathbb{Z}_+$ belongs to A or not. In other words, we want to determine the value of the characteristic function of A at the point x. Thus, we want to consider Towers of Algorithms for χ_A where the functions/relations at the lowest level shall be computable, and we again ask for the minimal height. More precisely, we consider

- the primary set $\Omega := \mathbb{Z}_+$,
- the evaluation set $\Lambda = \{\lambda\}$ consisting of the function $\lambda : \mathbb{Z}_+ \to \mathbb{C},\ x \mapsto x$,
- the metric space $\mathcal{M} := (\{\textit{true}, \textit{false}\}, d_{discr}) = (\{\textit{Yes}, \textit{No}\}, d_{discr})$,

where d_{discr} denotes the discrete metric, and consider all functions $\Xi : \Omega \to \mathcal{M}$ in the Arithmetical Hierarchy. In honour of Kleene and Shoenfield we call a Tower of Algorithms that is computable a Kleene-Shoenfield tower.

Definition 7.8 (Kleene-Shoenfield tower). *A tower of algorithms given by a family* $\{\Gamma_{n_k,\dots,n_1} : \Omega \to \mathcal{M} : n_k, \dots, n_1 \in \mathbb{N}\}$ *of functions at the lowest level is said to be a Kleene-Shoenfield tower, if the function*

$$\mathbb{N}^k \times \Omega \to \mathcal{M}, \quad (n_k, \dots, n_1, x) \mapsto \Gamma_{n_k,\dots,n_1}(x)$$

is computable. Given a computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ *as above, we will write* $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{KS}}$ *to denote the SCI with respect to a Kleene-Shoenfield tower.*

We can now present the main theorem, linking the SCI and the Arithmetical Hierarchy.

Theorem 7.9 (The SCI and the Arithmetical Hierarchy). *If* Ξ *is* Δ_{m+1} *then there exists a Kleene-Shoenfield tower of algorithms of height* m*. Conversely, if* $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{KS}} = m$ *then* Ξ *is* Δ_{m+1}*, but not* Δ_m*.*

This theorem has an immediate corollary that shows how the SCI can become arbitrarily large. In particular, for any $k \in \mathbb{N}$ there exists a problem that has SCI equal to k.

Corollary 7.10 (The SCI can become arbitrarily large). *For every* $k \in \mathbb{N}$ *there exists a problem function* Ξ *on* Ω *with* $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{KS}} = k$*.*

Theorem 7.9 follows immediately from a result by Shoenfield (Theorem 7.11), which has roots in the papers of Gold [43], Putnam [76], Shoenfield [88], and was discussed in [28, 81]. It is this result that builds the bridge between the SCI and the Arithmetical Hierarchy:

Theorem 7.11 (Shoenfield 1959, [72] (IV.1.19)). *For* $m \in \mathbb{N}$ *a function* $f : \mathbb{Z}_+ \to \{\textit{true}, \textit{false}\}$ *is* Δ_{m+1} *if and only if there is a computable function* $g : \mathbb{Z}_+^{m+1} \to \{\textit{true}, \textit{false}\}$ *such that*

$$f(y) = \lim_{x_1 \to \infty} \cdots \lim_{x_m \to \infty} g(y, x_1, \dots, x_m).$$

Moreover, Corollary 7.10 follows directly from Theorem 7.7. Next, we want to make the latter observations more precise, and we want to demonstrate how to transform a given tower into prenex normal form on the one hand, and on the other hand, how to define a Kleene-Shoenfield tower of minimal height for a given Δ_{m+1} relation. Actually, we do not do that for the present setting and for Kleene-Shoenfield towers, since

the proof can be found in the literature, but for our initial more general setting of decision making problems in Section 7.1.

7.5. **Alternating quantifier forms for General Towers.** We return to the more general setting of decision making problems of Section 7.1, with arbitrary Ω, a certain evaluation set Λ and $\mathcal{M} = \{\textit{true}, \textit{false}\} = \{\textit{Yes}, \textit{No}\}$. Inspired by the observations on the Arithmetical Hierarchy, we make the following definition:

Definition 7.12 (Alternating quantifier forms). *Given the general setup above we define the following:*

(i) *We say that $\Xi : \Omega \to \mathcal{M}$ permits a representation by an alternating quantifier form of length m if*

$$\Xi = (Q_m n_m) \cdots (Q_1 n_1) \Gamma_{n_m,\ldots,n_1},$$

where (Q_i) is a list of alternating quantifiers $(\forall)$ and $(\exists)$, and all $\Gamma_{n_m,\ldots,n_1} : \Omega \to \mathcal{M}$ are general algorithms in the sense of Definition 2.3.

(ii) *We say that Ξ is Σ_m if an alternating quantifier form of length m exists with Q_m being $(\exists)$, and that Ξ is Π_m if an alternating quantifier form of length m exists with Q_m being $(\forall)$.*

(iii) *We say that Ξ is Δ_m if Ξ is Σ_m and Π_m.*

The following theorem demonstrates how the SCI framework can be viewed as a generalization of the Arithmetical Hierarchy to arbitrary computational problems. In particular, one can define a hierarchy for any kind of tower. Here we do this for a general tower, and obviously, this can be done for any tower.

Theorem 7.13 (General Hierarchy). *Following Definition 7.12, the following is true.*

(i) *If $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{G}} \leq m$ then Ξ is Δ_{m+1}.*

(ii) *If Ξ is Σ_m or Π_m then $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{G}} \leq m$.*

(iii) *For $m \in \mathbb{N}$ we have that $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{G}} = m$ if and only if m is the smallest number with Ξ being Δ_{m+1}.*

We will call the hierarchy described above a *General Hierarchy*.

8. Roots of polynomials and Doyle-McMullen towers

In this section we recall the definition of a tower of algorithms from [38]. We will name this type of tower a Doyle-McMullen tower and demonstrate how the results in [67] and [38] can be put in a framework of the SCI. In particular, we will demonstrate how the construction of the Doyle-McMullen tower in [38] can be viewed as a tower of algorithms defined in Definition 2.4.

As mentioned in the introduction, one can compute zeros of a polynomial if one allows arithmetic operations and radicals and can pass to a limit. However, what if one cannot use radicals, but rather iterations of a rational map? A natural choice of such a rational map would be Newton's method. The only problem is that the iteration may not converge, and that motivated the question by Smale quoted in the introduction.

As we now know from [67] the answer is no, however, the results in [38] show that the quartic and the quintic can be solved with several rational maps and limits while this is not the case for higher degree polynomials. Below we first quote their results and then specify a particular tower of height three in the form that it can be viewed as a tower of algorithm in the sense of this paper.

8.1. **Doyle-McMullen towers.** A *purely iterative algorithm* [93] is a rational map [2]

$$T : \mathbb{P}_d \to \mathrm{Rat}_m,\ p \mapsto T_p$$

[2] I.e. it's a rational map of the coefficients of p.

which sends any polynomial p of degree $\leq d$ to a rational function T_p of a certain degree m. An important example of a purely iterative algorithm is *Newton's method*. Furthermore, Doyle and McMullen call a purely iterative algorithm *generally convergent* if

$$\lim_{n\to\infty} T_p^n(z) \text{ exists for } (p, z) \text{ in an open dense subset of } \mathbb{P}_d \times \hat{\mathbb{C}}.$$

Here $T_p^n(z)$ denotes the nth iterate $T_p^n(z) = T_p(T_p^{n-1}(z))$ of T_p. For instance, Newton's method is generally convergent *only* when $d = 2$. However, given a cubic polynomial $p \in \mathbb{P}_3$ one can define an appropriate rational function $q \in \mathrm{Rat}_3$ whose roots coincide with the roots of p, and for which Newton's method *is* generally convergent (see [67], Proposition 1.2). In [38] the authors provide a definition of a tower of algorithms, which we quote verbatim:

Definition 8.1 (Doyle-McMullen tower). *A tower of algorithms is a finite sequence of generally convergent algorithms, linked together serially, so the output of one or more can be used to compute the input to the next. The final output of the tower is a single number, computed rationally from the original input and the outputs of the intermediate generally convergent algorithms.*

Theorem 8.2 (McMullen [67]; Doyle and McMullen [38]). *For $\mathbb{P}_d$ there exists a generally convergent algorithm only for $d \leq 3$. Towers of algorithms exist additionally for $d = 4$ and $d = 5$ but not for $d \geq 6$.*

Note that, as shown in [89], there are generally convergent algorithms if one in addition allows the operation of complex conjugation. In the following we present how the Doyle-McMullen towers can be recast in the form of a general tower as defined in Definition 2.4.

8.2. **A height 3 tower for the quartic.** In the following $X, Y, \ldots$ denote variables in the polynomials while $x, y, \dots \in \mathbb{C}$. We build the tower following the standard reduction path, see e.g. [35]. Given

$$p(X) := X^4 + a_1X^3 + a_2X^2 + a_3X + a_4$$

one first transforms the equation by change of variable $Y = X + a_1/4$ to arrive into the polynomial

$$q(Y) := Y^4 + b_2Y^2 + b_3Y + b_4,$$

which one writes, with help of a parameter z, as $q(Y) = (Y^2 + z)^2 - r(Y, z)$ where

$$r(Y, z) = (2z - b_2)Y^2 - b_3Y + z^2 - b_4.$$

Here one wants a value of z such that $r(Y, z)$ becomes a square which requires the discriminant to vanish: $4(2z-b_2)(z^2-b_4)-b_3^2 = 0$. Viewing this as polynomial in Z, making a change of variable $W = Z+(1/6)b_2$ and scaling the polynomial to monic we arrive at asking for a root of

$$s(W) := W^3 + c_2W + c_3. \tag{8.1}$$

As all these are rational computations on the coefficients of p, we shall not express them explicitly.

We denote by $N(f, \xi_0)$ the function in Newton's iteration with initial value ξ_0:

$$\xi_{j+1} := N(f(\xi_j)) \ \text{ where } \ N(f(\xi)) = \xi - \frac{f(\xi)}{f'(\xi)}$$

and further by N_j the mapping from initial data to the j^{th} iterate $N_j : (f, \xi_0) \mapsto \xi_j$. We shall apply Newton's iteration to the rational function [38]

$$t(W) := \frac{s(W)}{3c_2W^2 + 9c_3W - c_2^2}.$$

Thus $w_j = N_j(t, w_0)$ denotes the j^{th} iterate w_j for a zero for $s(w) = 0$. This iteration converges in an open dense set of initial data. Denote $w_\infty := \lim_{j\to\infty} w_j$. Now we change the variable $Z = W - (1/6)b_2$ and, denoting by z_j and z_∞ the corresponding values, we obtain $r(Y, z_\infty)$ as a square:

$$r(Y, z_\infty) = (2z_\infty - b_2)\left(Y - \frac{b_3}{2(2z_\infty - b_2)}\right)^2.$$

To find a zero of $q(Y)$ we shall need to have a generally convergent iteration for $\sqrt{2z - b_2}$. Thus, we set $u_j(V) := V^2 + b_2 - 2z_j$ and apply Newton's method for this, starting with initial guess v_0 and iterating k times and set $v_{k,j} := N_k(u_j, v_0)$. From $q(Y) = (Y^2 + z_\infty)^2 - r(Y, z_\infty) = 0$ we move to solve one of the factors

$$Q(Y) = Y^2 + z_\infty - \sqrt{2z_\infty - b_2}\left(Y - \frac{b_3}{2(2z_\infty - b_2)}\right) = 0.$$

However, we can do this only based on approximative values for the parameters, so we set

$$Q_{k,j}(Y) = Y^2 + z_j - v_{k,j}\left(Y - \frac{b_3}{2(2z_j - b_2)}\right) = 0.$$

Now apply Newton's iteration to this, say n times, using starting value y_0 and denote the output by $y_{n,k,j}$:

$$y_{n,k,j} = N_n(Q_{k,j}, y_0).$$

Finally, we set $x_{n,k,j} = y_{n,k,j} - a_1/4$ in order to get an approximation to a root of p. Suppose now $j = n_1, k = n_2, n = n_3$. If $n_1 \to \infty$ then $w_{n_1} \to w_\infty$ and hence $z_{n_1} \to z_\infty$, too. It is natural to denote $u(V) := V^2 + b_2 - 2z_\infty$ and correspondingly $v_{n_2} := N_{n_2}(u, v_0)$ and

$$Q_{n_2}(Y) = Y^2 + z_\infty - v_{n_2}\left(Y - \frac{b_3}{2(2z_\infty - b_2)}\right) = 0.$$

Then in an obvious manner $x_{n_3,n_2} = N_{n_3}(Q_{n_2}, y_0) - a_1/4$. Then we have $\lim_{n_1\to\infty} x_{n_3,n_2,n_1} = x_{n_3,n_2}$. If we denote $x_{n_3} = N_{n_3}(Q, y_0) - a_1/4$, then clearly $\lim_{n_2\to\infty} x_{n_3,n_2} = x_{n_3}$. Finally $x_\infty = \lim_{n_3\to\infty} x_{n_3}$ is a root of p.

The link to the SCI. One special feature of these towers which are build on generally convergent algorithms is the following: in addition to the polynomial p, the initial values for the iterations have to be read into the process via evaluation functions. Denoting the initial values for the three different Newton's iterations by $d_0 = (w_0, v_0, y_0) \in \mathbb{C}^3$ we can now put this Doyle-McMullen tower in the form of a general tower as defined in Definition 2.4, with the slight weakening that, for each $p \in \mathbb{P}_4$, the tower might converge only at a dense subset of initial values. In particular, set

$$\Gamma_{n_3} : \mathbb{P}_4 \times \mathbb{C}^3 \to \mathbb{C}, \text{ by } (p, d_0) \mapsto x_{n_3},$$
$$\Gamma_{n_3,n_2} : \mathbb{P}_4 \times \mathbb{C}^3 \to \mathbb{C} \text{ by } (p, d_0) \mapsto x_{n_3,n_2},$$
$$\Gamma_{n_3,n_2,n_1} : \mathbb{P}_4 \times \mathbb{C}^3 \to \mathbb{C} \text{ by } (p, d_0) \mapsto x_{n_3,n_2,n_1}.$$

Thus, if we let $\Omega = \mathbb{P}_4 \times \mathbb{C}^3$ and $\Xi, \mathcal{M}$ be as in Example 2.1 (III), and complement Λ by the mappings that read w_0, v_0, y_0 from the input, then by the construction above and Theorem 8.2 we have that

$$\mathrm{SCI}(\Xi, \Omega)_{\mathrm{DM}} \in \{2, 3\}.$$

8.3. **A height 3 tower for the quintic.** Let

$$p(X) = X^5 + a_1X^4 + a_2X^3 + a_3X^2 + a_4X + a_5$$

be the given quintic. Doyle and McMullen [38] give a generally convergent algorithm for the quintic in Brioschi form. Thus, one needs first to bring the general quintic to Brioschi form, then apply the iteration and finally construct at least one root for $p(X)$. In the following we outline a path for doing this, which follows L. Kiepert [58] except that the Brioschi quintic is solved by Doyle-McMullen iteration rather than by using Jacobi sextic. This path can be found in [59].

One begins applying a Tschirnhaus transformation $Y = X^2 - uX + v$ to arrive into *principal* form

$$q(Y) = Y^5 + b_3Y^2 + b_4Y + b_5.$$

Here v is obtained from a linear equation but to solve u one needs to solve a quadratic equation $Q(U) = U^2 + \alpha U + \beta$, where the coefficients α, β are rational expressions of the coefficients of $p(X)$, (see for example p. 100, eq. (6.2-9) in [59]).

Here is the first application of Newton's method. We are given an initial value u_0 and iterate j times $u_j = N_j(Q, u_0)$. We may assume that v is known exactly but we only have an approximation u_j to make the transformation. So, suppose the Newton iteration converges to u_∞. Thus, we make the transformation using u_j and *force* the coefficients $b_{2,j} = b_{1,j} = 0$ while keep the others as they appear. The transformation being continuous yields polynomials

$$q_j(Y) = Y^5 + b_{3,j}Y^2 + b_{4,j}Y + b_{5,j},$$

whose roots shall converge to those of $q(Y)$. The next step is to transform $q_j(Y)$ into Brioschi form. Let the Brioschi form corresponding to the exact polynomial $q(Y)$ be denoted by $B(Z)$

$$B(Z) = Z^5 - 10CZ^3 + 45C^2Z - C^2 = 0, \tag{8.2}$$

while with $B_j(Z)$ we denote the exact Brioschi form corresponding to $q_j(Y)$. The transformation from $q(Y)$ to $B(Z)$ is of the form

$$Y = \frac{\lambda + \mu Z}{(Z^2/C) - 3}. \tag{8.3}$$

Here λ satisfies a quadratic equation with coefficients being polynomials of the coefficients in the principal form (p. 107, eq. (6.3-28) in [59]). Let us denote that quadratic by $R(L)$ when it comes from $q(Y)$ and by $R_j(L)$ when it comes from $q_j(Y)$ respectively. Thus here we meet our second application of Newton's method. So, we denote by

$$\lambda_{k,j} := N_k(R_j, \lambda_0)$$

the output of iterating k times for a solution of $R_j(L) = 0$. And, in a natural manner, we denote also

$$\lambda_k = N_k(R, \lambda_0) \quad \text{and} \quad \lambda = \lim_{k\to\infty} N_k(R, \lambda_0).$$

The corresponding values of $\mu_{k,j}, \mu_k$ and μ are then obtained by simple substitution (p. 107, eq. (6.3-30) in [59]). The Tschirnhaus transformation with exact values (λ, μ) transforms the equation not yet to the Brioschi form with just one parameter C but such that the constant term may be different. However, the last step is just a simple scaling and then one is in the Brioschi form (8.2). However, when we apply the transformation with the approximated values $(\lambda_{k,j}, \mu_{k,j})$ or with (λ_k, μ_k) we do not arrive at the Brioschi form. So, we *force* the coefficients of the fourth and second powers to vanish and replace the coefficients of the first power to match with the coefficients in the third power. Finally, after scaling the constant terms we have the Brioschi quintics $B_{k,j}$ and B_k, e.g.

$$B_{k,j}(Z) = Z^5 - 10C_{k,j}Z^3 + 45C_{k,j}^2Z - C_{k,j}^2 = 0. \tag{8.4}$$

Provided that the Newton iterations converge, that is, the initial values (u_0, λ_0) are generic, these quintics converge to the exact one.

Here we apply the generally convergent iteration by Doyle and McMullen [38]. They specify a rational function

$$T_C(Z) = z - 12\frac{g_C(Z)}{g'_C(Z)}$$

where g is a polynomial of degree 6 in the variable C and of degree 12 in Z. Starting from an initial guess w_o from an initial guess $w_{n+1} = T_C(T_C(w_n))$ to convergence and applying T_C still once, we obtain, after a finite rational computation with these two numbers, two roots of the Brioschi, say z_I and z_{II}. If applied

to the approximative quintics and if the iteration is truncated after n steps, together with the corresponding postprocessing, we have obtained e.g. a pair $(z_{I,n,k,j}, z_{II,n,k,j})$.

What remains is to invert the Tschirnhaus transformations. Suppose z is a root of the exact Brioschi form (8.2). Then the corresponding root of the principal quintic is obtained immediately from (8.3)

$$ty = \frac{\lambda + \mu z}{(z^2/C) - 3}.$$

Naturally, we can only apply this using approximated values for the parameters. Finally, one needs to transform the (approximative roots) of the principal quintic to (approximative) roots for the original general quintic $p(X)$. This is done by a rational function $X = r(Y)$ where $r(Y)$ is of second order in Y and the coefficients are polynomials of the coefficients if the original $p(X)$ and u and v (p. 127, eq. (6.8-3) in [59]). Again, we would be using only approximative values u_j in place of the exact u. In any case, at the end we obtain a pair of approximations to the roots of the original quintic. If we put $n_1 = j, n_2 = k$ and $n_3 = n$, then this pair could be denoted by $(x_{I,n_3,n_2,n_1}, x_{II,n_3,n_2,n_1})$.

The link to the SCI. In the same way as with the quartic, we assume that the initial value $d_0 = (u_0, \lambda_0, w_0) \in \mathbb{C}^3$ is generic, so that all iterations converge for large enough values and since the transformations are continuous functions of the parameters in it, all necessary limits exist and match with each others. The functions Γ_{n_3,n_2,n_1} can then be identified in a natural manner:

$$\begin{aligned}
\Gamma_{n_3} &: \mathbb{P}_5 \times \mathbb{C}^3 \to \mathbb{C}^2, \text{ by } (p, d_0) \mapsto (x_{I,n_3}, x_{II,n_3}),\\
\Gamma_{n_3,n_2} &: \mathbb{P}_5 \times \mathbb{C}^3 \to \mathbb{C}^2 \text{ by } (p, d_0) \mapsto (x_{I,n_3,n_2}, x_{II,n_3,n_2}),\\
\Gamma_{n_3,n_2,n_1} &: \mathbb{P}_5 \times \mathbb{C}^3 \to \mathbb{C}^2 \text{ by } (p, d_0) \mapsto (x_{I,n_3,n_2,n_1}, x_{II,n_3,n_2,n_1}),
\end{aligned}$$

where $(x_{I,n_3,n_2}, x_{II,n_3,n_2})$ and (x_{I,n_3}, x_{II,n_3}) are the limits as $n_1 \to \infty$ and $n_2 \to \infty$ respectively. These limits exist for initial values in an open dense subset of $\mathbb{C}^3$. Hence, we let $\Omega = \mathbb{P}_5 \times \mathbb{C}^3$, and $\Xi, \mathcal{M}, \Lambda$ be as in case of the quartic. Then, by the construction above and Theorem 8.2 we have, again in a slightly weakened sense, that

$$\mathrm{SCI}(\Xi, \Omega)_{\mathrm{DM}} \in \{2, 3\}.$$

8.4. **Particular initial guesses and height one towers.** The special feature of the above mentioned Doyle-McMullen towers is that they address the question whether one can achieve converge to the roots of a polynomial p for (almost) arbitrary initial guesses. With slight change of perspective one might also ask the question how big the SCI gets if one applies purely iterative algorithms *after a suitable clever choice* of initial values. And indeed, the answer to that question is really satisfactory: For polynomials of arbitrary degree one can compute the whole set of roots (more precisely: approximate it in the sense of the Hausdorff distance) by a tower of height one which just consists of Newtons method.

The key tool for the choice of the initial values is the main theorem of [54]:

Theorem 8.3 (Hubbard, Schleicher and Sutherland [54]). *For every $d \geq 2$ there is a set S_d consisting of at most $1.11 d \log^2 d$ points in $\mathbb{C}$ with the property that for every polynomial p of degree d and every root z of p there is a point $s \in S_d$ such that the sequence of Newton iterates $\{s_n\}_{n\in\mathbb{N}} := \{N_p^n(s)\}_{n\in\mathbb{N}}$ converges to z. In particular, the proof is constructive, and these sets S_d can easily be computed.*

A further important property of Newton's method is that, in case of converge, the speed is at least linear: If $z_n := N_p^n(s)$ tend to a root z of p then there exists a constant c such that $|z_n - z| \leq c/n$. Finally we have the following.

Proposition 8.4. *Let p be a polynomial of degree d, $\epsilon > 0$ and $z_n := N_p^n(s)$. If $|z_n - z_{n+1}| < \frac{\epsilon}{d}$ then there is a root z of p with $|z_n - z| < \epsilon$.*

Proof. We have $\left|\frac{p(z_n)}{p'(z_n)}\right| = |z_n - z_{n+1}| < \frac{\epsilon}{d}$, hence $|p(z_n)| < \frac{\epsilon|p'(z_n)|}{d}$. Decompose $p(x) = a\Pi_{i=1}^d(x-x_i)$, notice that $p'(x) = a\sum_{j=1}^d \Pi_{i=1,i\neq j}^d(x-x_i)$, choose j such that $|\Pi_{i=1,i\neq j}^d(z_n - x_i)|$ is maximal, and conclude that

$$|a\Pi_{i=1}^d(z_n - x_i)| = |p(z_n)| < \frac{\epsilon|p'(z_n)|}{d} \leq \epsilon|a\Pi_{i=1,i\neq j}^d(z_n - x_i)|,$$

thus $|z_n - x_j| < \epsilon$. Now $z = x_j$ is a root as asserted. □

Let p be a polynomial of degree d. For each $s \in S_d$ let s_n denote the nth Newton iterates of s, and define

$$\Gamma_n(p) := \left\{ s_n : s \in S_d, |s_n - s_{n+1}| < \frac{1}{\sqrt{n}} \right\}. \tag{8.5}$$

Then $(\Gamma_n(p))$ converges to the set $\Sigma(p)$ of all zeros of p in the Hausdorff metric. Indeed, let z be a zero of p. By Theorem 8.3 there is an initial value $s \in S_d$ such that $s_n = N_p^n(s)$ tend to z with at least linear speed, i.e.

$$|s_n - s_{n+1}| \leq |s_n - z| + |s_{n+1} - z| \leq \frac{2c}{n} < \frac{1}{\sqrt{n}}$$

for all large n, hence $s_n \in \Gamma_n(p)$ for all large n. Conversely, each $s_n \in \Gamma_n(p)$ has the property that its distance to the set $\Sigma(p)$ is less than $\epsilon = \frac{d}{\sqrt{n}}$ by Proposition 8.4.

Therefore we define $\Omega_d = \mathbb{P}_d$ to be the set of polynomials of degree d, $\mathcal{M}$ the set of finite subsets of $\mathbb{C}$ equipped with the Hausdorff metric, and $\Xi : \Omega_d \to \mathcal{M}$ be the mapping that sends $p \in \Omega_d$ to the set of its zeros. Further Λ_d shall consist of the evaluation functions that read the coefficients of the polynomial $p \in \Omega_d$, and the constant functions with the values $s \in S_d$. Note again that these values can be effectively constructed.

Theorem 8.5. *Consider* $(\Xi, \Omega_d, \mathcal{M}, \Lambda_d)$ *as above. Then the algorithms* (8.5) *define an arithmetic tower of height one for the computation of the roots of each input polynomial* p, *thus* $\mathrm{SCI}(\Xi, \Omega_d, \mathcal{M}, \Lambda_d)_{\mathrm{A}} \leq 1$. *Moreover, this tower employs just Newton's Method, i.e. a purely iterative algorithm.*

9. Open Problems

Establishing the SCI opens up for a long list of open problems. In particular, we now need to classify all types of computational problems in terms of the SCI. One could think of essentially four main categories of problems:

- **(I)** Problems with SCI equal to zero. This class contains most of the problems in classical complexity theory.
- **(IIa)** Problems with SCI equal to one, where one also has error control. This set includes integration problems, ODE's, root finding of polynomials, etc., and is the core of information based complexity [96] as well as parts of real number complexity theory [11].
- **(IIb)** Problems with SCI equal to one, but where there is no error control. This set includes for example problems of computing spectra of tridiagonal self-adjoint infinite matrices like real discrete Schrödinger operators.
- **(III)** Problems with SCI greater than one. This class includes the main problems discussed in this paper, and this is the class that needs most work. In particular, we need a full classification theory for all problems in this class determining the SCI. What is presented in this paper is just the beginning. We predict that this class is vast and suggestions to problems that may be in this class follow below.

Remark 9.1 (Adding structure to reduce the SCI). Note that classifying subclasses of Class III according to the SCI is about determining what kind of extra structure and information needs to be added to reduce the SCI from k to $k-1$ and $k-2$ and so forth. This means also that one has to invent new algorithms (or towers of algorithms) as the tower used to obtain $\mathrm{SCI} = k$ obviously cannot be used for problems with $\mathrm{SCI} = k-1$. Thus, classifying all subclasses of III will lead to vast numbers of new algorithms.

9.1. **Potential problems with SCI greater than one.**

9.1.1. *Quantum mechanics.* It is indeed likely that some of the fundamental computational quantum mechanics have SCI greater than one. We have already discussed upper bounds of the SCI for computing spectra of Schrödinger operators

$$H = -\Delta + V, \qquad V : \mathbb{R}^d \to \mathbb{C}, \tag{9.1}$$

and we believe that the SCI is equal to two for computing spectra of these operators when considering the setup in Section 4 where V is bounded and we are using an Arithmetic tower.

Similarly, this may very well be the case for Dirac operators. In particular, consider the same computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ as in Section 4, however, with the Schrödinger operator replaced by the Dirac operator. More precisely, let $\mathcal{H} = \oplus_{k=1}^4 L^2(\mathbb{R}^3)$ and define (formally) $\tilde{P}_j$ on $\mathcal{H}$ by $\tilde{P}_j = \oplus_{k=1}^4 P_j$, with $P_j = -i\frac{\partial}{\partial x_j}$, for $j = 1, 2, 3$, where P_j is formally defined on $L^2(\mathbb{R}^3)$. Let $H_0 = \sum_{j=1}^3 \alpha_j \tilde{P}_j + \beta$, where α_j and β are 4-by-4 matrices satisfying the commutation relation

$$\alpha_j\alpha_k + \alpha_k\alpha_j = 2\delta_{jk}I, \qquad j, k = 1, 2, 3, 4, \qquad \alpha_4 = \beta.$$

Then it is well known that H_0 is self-adjoint on $\oplus_{k=1}^4 W_{2,1}(\mathbb{R}^3)$ where $W_{2,1}(\mathbb{R}^3) = \{f \in L^2(\mathbb{R}^3) : \mathcal{F}f \in L_1^2(\mathbb{R}^3)\}$ and $L_1^2(\mathbb{R}^3) = \{f \in L^2(\mathbb{R}^3) : (1 + |\cdot|^2)^{1/2} f \in L^2(\mathbb{R}^3)\}$. Let $V \in L^\infty(\mathbb{R}^3)$ and define the Dirac operator

$$H_D = H_0 + \bigoplus_{k=1}^4 V, \qquad \mathcal{D}(H) = \bigoplus_{k=1}^4 W_{2,1}(\mathbb{R}^3).$$

The question is: what is the SCI of computing spectra of such Dirac operators? Note that the similarity between the Dirac operator and the Schrödinger operator suggest that we may very well be able to use the techniques from this paper to get the upper bounds.

9.1.2. *Infinite dimensional optimization and inverse problems.* Given an $A \in \mathcal{B}(l^2(\mathbb{N}))$ and $y \in l^2(\mathbb{N})$ and the optimization problem of finding

$$x \in \operatorname*{argmin}_{\eta \in l^p(\mathbb{N})} \|\eta\|_{l^p} \text{ subject to } \|A\eta - y\| \leq \delta, \qquad \delta \geq 0, \quad p \in [1, \infty), \tag{9.2}$$

where we assume that this problem is feasible for the given δ, meaning that there exists at least one minimizer. Such problems are popular in sampling theory and compressed sensing [1, 2, 25, 37]. In this case we may consider the computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$ similar to what we have in Section 5, however, we would first consider the closed metric space $\{\eta : \|A\eta - y\| \leq \delta\}$ and let $\mathcal{M}$ be the metric quotient space when identifying all the minimizer. Also, $\Xi = [x]$ where x is a minimizer and $[x]$ denotes the equivalence class corresponding to x. Note that if $\Omega = \Omega_1 \times \Omega_2$ where $\Omega_1 = \{A : A = A^*, A \text{ is compact}, \|A\| \leq M\}$ for some $M > 0$ and $\Omega_2 = \{y : y \in \mathrm{Ran}(A)\}$, where $\mathrm{Ran}(A)$ denotes the range of A, then it can be shown [50] that $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{G}} \geq 2$. This suggests that even in the case where there is a substantial amount of extra structure, there are many infinite-dimensional optimization problems and inverse problems with $\mathrm{SCI} > 1$.

Interestingly, if we instead consider the problem of finding

$$x \in \operatorname*{argmin}_{\eta \in l^p(\mathbb{N})} \|A\eta - y\|^2 + \lambda\|\eta\|_{l^p}, \qquad \lambda > 0, \quad p \in [1, 2],$$

with the similar setup as above then it can be shown [50] that $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{G}} = 1$. Similarly, we may consider the problems of finding

$$g \in \operatorname*{argmin}_{u \in \mathrm{BV}(\tilde{\Omega})} \frac{1}{2}\int_{\tilde{\Omega}} (Tu - f)^2\, dx + \lambda \mathrm{TV}(u), \qquad \lambda > 0, \tag{9.3}$$

and

$$g \in \operatorname*{argmin}_{u \in \mathrm{BV}(\tilde{\Omega})} \mathrm{TV}(u) \text{ subject to } \int_{\tilde{\Omega}} (Tu - f)^2 \, dx \leq \delta, \qquad \lambda > 0, \tag{9.4}$$

where $\tilde{\Omega} \subset \mathbb{R}^d$ is some appropriate domain, $T : L^2(\tilde{\Omega}) \to L^2(\mathbb{R}^d)$ is some linear operator, f is some function, possibly not even in the range of T, $\mathrm{BV}(\tilde{\Omega})$ denotes the set of functions that have bounded variations and

$$\mathrm{TV}(u) = \sup \left\{ \int_{\tilde{\Omega}} u \operatorname{div} u \, dx : \, v \in C_c^1(\tilde{\Omega}, \mathbb{R}^2), \, \|v\|_\infty \leq 1 \right\}$$

where $C_c^1(\tilde{\Omega}, \mathbb{R}^2)$ is the set of continuously differentiable vector functions of compact support contained in $\tilde{\Omega}$. The optimization problem (9.3) is highly popular in imaging [80] for example and a well known regularization technique in inverse problems. The fact that solving (9.2) has SCI greater than one indeed suggests that solving (9.4) even when T is compact also has SCI greater than one.

These are just two popular examples, however, there are essentially countless number of infinite-dimensional optimization problems [40] of the same nature as the ones above, thus we predict that there are many of those problems with $\mathrm{SCI} > 1$.

9.1.3. *Operator semigroups and PDEs in unbounded domains.* Suppose that A is a closed linear operator which is bounded from below: $\langle Ax, x \rangle \geq \mu \|x\|^2$ for some $\mu \in \mathbb{R}$ and $x \in \mathcal{D}(A)$, and one is to compute, for a fixed $t > 0$, e^{-tA} or $e^{-tA}x_0$ for some x_0 in the Hilbert space. In the former case the task is to compute approximations which converge in the operator norm and in the latter in the norm of the underlying space. The semigroup generated by the Schrödinger operator $-iH$ is naturally of particular interest: given $H = -\Delta + V$ compute

$$e^{-itH}\psi, \qquad \psi \in L^2(\mathbb{R}^n), \qquad t > 0. \tag{9.5}$$

If we consider ψ to be known, then solving (9.5) can be done by solving a PDE, namely the time dependent Schrödinger equation. However, as V is defined on the whole space, this PDE would not have any boundary.

A typical method for computing solutions to PDEs in unbounded domains is to impose non-physical boundaries with some boundary conditions on them, so that the solution within is close (in some sense) to the correct solution of the original problem. This arises is *wave*-type equations, including the (time-dependent) Schrödinger equation. A notable result in this field is that of Engquist and Majda [41], where a method for imposing artificial boundary conditions that minimize non-physical reflections is introduced.

However, given a type of PDEs with an unbounded domain, this still leaves the question of *what is the associated* SCI unanswered. More precisely, one could ask for computational solutions that converge to the true solution of a given problem on some finite time interval $[0, t]$. Is this possible in one limit? Note that the method of adding artificial boundary conditions, if such method is successful, would essentially have two limiting procedures. In particular, one limit for each boundary chosen, and one limit when the boundary tends to infinity. Therefore, it is not unreasonable to believe that the $\mathrm{SCI} > 1$ for certain types of PDEs in unbounded domains.

9.1.4. *Computational harmonic analysis.* A problem in computational harmonic analysis is to compute frame bounds [26] for a given frame for $L^2(R^d)$. As this means computing extremal spectral values of a self-adjoint infinite matrix we are faced with a several limit problem. First, computing spectra of a self-adjoint operator has $\mathrm{SCI} = 2$. Second, producing matrix elements from the frame elements $f \in L^2(R^d)$ may indeed add an extra limit, unless this limit can be collapsed with one of the other limits. However, this question is completely open and computing frame bounds may have $\mathrm{SCI} > 1$.

10. Proofs of theorems in Section 3

We start the sections on the proofs of our main results with a simple but fundamental observation on the smallest singular values $\sigma_1(B)$ of matrices $B \in \mathbb{C}^{m\times n}$, which constitutes one of the corner stones for most of the general algorithms we will construct in the subsequent proofs.

Proposition 10.1. *Given a matrix $B \in \mathbb{C}^{m\times n}$ and a number $\epsilon > 0$ one can test with finitely many arithmetic operations of the entries of B whether the smallest singular value $\sigma_1(B)$ of B is greater than ϵ.*

Proof. The matrix B^*B is self-adjoint and positive semidefinite, hence has its eigenvalues in $[0,\infty)$. The singular values of B are the square roots of these eigenvalues of B^*B. The smallest singular value is greater than ϵ iff the smallest eigenvalue of B^*B is greater than ϵ^2, which is the case iff $C := B^*B - \epsilon^2 I$ is positive definite. It is well known that C is positive definite if and only if the pivots left after Gaussian elimination (without row exchange) are all positive. Thus, if C is positive definite, Gaussian elimination leads to pivots that are all positive, and this requires finitely many arithmetic operations. If C is not positive definite, then at some point a pivot is zero or negative, at this point the algorithm aborts. □

Remark 10.2. In practice it may be advisable to use the Cholesky decomposition when determining if C is positive definite or not. This is simply for stability purposes as discussed in [50]. However, introducing the Cholesky decomposition requires evaluation of a radical. As Proposition 10.1 is the mainstay in most Arithmetic towers constructed in this paper, one can simply replace the Gaussian elimination with the Cholesky decomposition and obtain a Radical tower instead that may be more appropriate for computations.

Proof of Theorem 3.3. Since $\Omega_1 \supset \Omega_2 \supset \Omega_3$ it obviously suffices to show that $\mathrm{SCI}(\Xi,\Omega_1)_{\mathrm{A}} \leq 2$ and $\mathrm{SCI}(\Xi,\Omega_3)_{\mathrm{G}} \geq 2$. We will start by showing that $\mathrm{SCI}(\Xi,\Omega_1)_{\mathrm{A}} \leq 2$, thus we will show the existence of an Arithmetic tower of algorithms of height 2 for $\Xi = \mathrm{sp}$ on the set Ω_1, where Ω_1 is determined by g in (3.2). So let $A \in \Omega_1$ and moreover let

$$\gamma(z) := \min\{\sigma_1(A - zI), \sigma_1(A^* - \overline{z}I)\} = \|(A - zI)^{-1}\|^{-1},$$

where $\sigma_1(B) := \inf\{\|B\xi\| : \xi \in l^2(\mathbb{N}), \|\xi\| = 1\}$ is known as the lower norm, the injection modulus, or the smallest singular value of $B \in \mathcal{B}(l^2(\mathbb{N}))$. To see why $\min\{\sigma_1(A - zI), \sigma_1(A^* - \overline{z}I)\} = \|(A - zI)^{-1}\|^{-1}$ see for example [49].

Step I (The construction of the tower of algorithms): Let $g : [0,\infty) \to [0,\infty)$ be as in the statement of Theorem 3.3, in particular, continuous, vanishing only at $x = 0$ and tending to ∞ as $x \to \infty$. Note that without loss of generality we can also assume that g is strictly increasing and $g(x) \leq x$ for all x. Then the inverse function $h(y) := g^{-1}(y) : [0,\infty) \to [0,\infty)$ is well defined, continuous, strictly increasing, $h(y) \geq y$ for every y, and $\lim_{y\to 0} h(y) = 0$.

Let $K \subset \mathbb{C}$ be a compact set and $\delta > 0$. We introduce a δ-grid for K by $G^\delta(K) := (K + B_\delta(0)) \cap (\delta(\mathbb{Z} + \mathrm{i}\mathbb{Z}))$, where $B_\delta(0)$ denotes the closed ball of radius δ centered at 0. Without loss of generality we may assume that δ^{-1} is an integer, and obviously, $G^\delta(K)$ is finite. Moreover, introduce $h_\delta(y) := \min\{k\delta : k \in \mathbb{N}, g(k\delta) > y\}$ and observe that for each y, evaluating $h_\delta(y)$ requires only finitely many evaluations of g. Also, notice that $h(y) \leq h_\delta(y) \leq h(y) + \delta$. For a given function $\zeta : \mathbb{C} \to [0,\infty)$ we define sets $\Upsilon_K^\delta(\zeta)$ as follows: For each $z \in G^\delta(K)$ let $I_z := B_{h_\delta(\zeta(z))}(z) \cap (\delta(\mathbb{Z} + \mathrm{i}\mathbb{Z}))$. Further

- If $\zeta(z) \leq 1$ then introduce the set M_z of all $w \in I_z$ for which $\zeta(w) \leq \zeta(v)$ holds for all $v \in I_z$.
- Otherwise, if $\zeta(z) > 1$, just set $M_z := \emptyset$.

Now define

$$\Upsilon_K^\delta(\zeta) := \bigcup_{z \in G^\delta(K)} M_z. \tag{10.1}$$

Notice that for the computation of $\Upsilon_K^\delta(\zeta)$ only finitely many evaluations of ζ and g are required. To define the lowest level of the tower by this construction we introduce appropriate functions

$$\zeta_{m,n}(z) := \min\{k/m : k \in \mathbb{N},\ k/m \geq \min\{\sigma_1(P_n(A - zI)P_m|_{\mathrm{Ran}(P_m)}), \sigma_1(P_n(A^* - \overline{z}I)P_m|_{\mathrm{Ran}(P_m)})\}\}. \tag{10.2}$$

Then we define

$$\begin{aligned} \Gamma_{m,n}(A) = \Gamma_{m,n}(\{A_{ij}\}_{i,j=1}^{n,m}) &:= \Upsilon_{B_m(0)}^{1/m}(\zeta_{m,n}), \\ \Gamma_m(A) &:= \lim_{n\to\infty} \Upsilon_{B_m(0)}^{1/m}(\zeta_{m,n}), \end{aligned} \tag{10.3}$$

where we will show that the limit exists. To show that this provides an arithmetic tower of algorithms for Ξ we firstly mention that each of the mappings $A \mapsto \Gamma_{m,n}(A)$ is a general algorithm in the sense of Definition 2.3. Moreover, as mentioned above, the computation of $\Upsilon_{B_m(0)}^{1/m}(\zeta_{m,n})$ requires only finitely many evaluations of $\zeta_{m,n}$, and the finite number of constants $g(k/m) \leq 1$, $k = 1, \ldots$. Hence it suffices to demonstrate that, for a single $z \in \mathbb{C}$, the evaluation of $\zeta_{m,n}(z)$ requires finitely many arithmetic operations of the evaluations A_f, $f \in \Lambda_{\Gamma_{m,n}}(A)$. This is done as follows. For $k \in \mathbb{N}$, we start with $k = 1$, then:

- Check whether $\min\{\sigma_1(P_n(A - zI)P_m|_{\mathrm{Ran}(P_m)}), \sigma_1(P_n(A^* - \overline{z}I)P_m|_{\mathrm{Ran}(P_m)})\} \leq k/m$.
- If not let $k = k + 1$ and repeat, otherwise $\zeta_{m,n}(x) = k/m$.

Note that the first step requires finitely many arithmetic operations of the complex numbers A_f, $f \in \Lambda_{\Gamma_{m,n}}(A)$ by Proposition 10.1, and the loop will clearly terminate for a finite k. Secondly, we have to show the desired convergence as $n \to \infty$ and then as $m \to \infty$.

Step II: $\Gamma_{m,n}(A) \to \Gamma_m(A)$, and $\Gamma_m(A) \to \Xi_1(A)$. To prove Step II we need the following result.

Claim: Let K be a compact set containing the spectrum of A and $0 < \delta < \epsilon < 1/2$. Further assume that ζ is a function with $\|\zeta - \gamma\|_{\infty,\hat{K}} := \|(\zeta - \gamma)\chi_{\hat{K}}\|_\infty < \epsilon$ on $\hat{K} := (K + B_{h(\mathrm{diam}(K)+2\epsilon)+\epsilon}(0))$, where $\chi_{\hat{K}}$ denotes the characteristic function of $\hat{K}$. Finally, let

$$u(\xi) := \max\{h(3\xi + h(t + \xi) - h(t)) + \xi : t \in [0, 1]\}. \tag{10.4}$$

Then we have that

$$d_{\mathrm{H}}(\Upsilon_K^\delta(\zeta), \mathrm{sp}(A)) \leq u(\epsilon) \text{ and } \lim_{\xi\to 0} u(\xi) = 0.$$

To prove the claim, let $z \in G^\delta(K)$ and notice that $I_z \subset \hat{K}$ since, for every $v \in I_z$,

$$\begin{aligned} |z - v| &\leq h_\delta(\zeta(z)) \leq h_\delta(\gamma(z) + \epsilon) \leq h(\mathrm{dist}(z, \mathrm{sp}(A)) + \epsilon) + \delta \\ &\leq h(\mathrm{diam}(K) + \delta + \epsilon) + \delta. \end{aligned} \tag{10.5}$$

Suppose that $M_z \neq \emptyset$. Note that by (3.2), the monotonicity of h, and the compactness of $\mathrm{sp}(A)$ there is a $y \in \mathrm{sp}(A)$ of minimal distance to z with $|z - y| \leq h(\gamma(z))$. Since $\|\zeta - \gamma\|_{\infty,\hat{K}} < \epsilon$ we get $|z - y| \leq h(\zeta(z) + \epsilon)$. Hence, at least one of the $v \in I_z$, let's say v_0, satisfies $|v_0 - y| < h(\zeta(z) + \epsilon) - h(\zeta(z)) + \delta$. Noting again that $\gamma(v_0) \leq \mathrm{dist}(v_0, \mathrm{sp}(A))$, we get $\zeta(v_0) < \gamma(v_0) + \epsilon < h(\zeta(z) + \epsilon) - h(\zeta(z)) + 2\epsilon$. By the definition of M_z this estimate now holds for all points $w \in M_z$ and we conclude that, for all $w \in M_z$,

$$\begin{aligned} \mathrm{dist}(w, \mathrm{sp}(A)) &= h(g(\mathrm{dist}(w, \mathrm{sp}(A)))) \leq h(\gamma(w)) \\ &\leq h(\zeta(w) + \epsilon) \leq h(h(\zeta(z) + \epsilon) - h(\zeta(z)) + 3\epsilon). \end{aligned} \tag{10.6}$$

This observation holds for every $z \in G^\delta(K)$ and all $w \in M_z$, hence all points in $\Upsilon_K^\delta(\zeta)$ are closer to $\mathrm{sp}(A)$ than $u(\epsilon)$.

Conversely, take any $y \in \mathrm{sp}(A) \subset K$. Then there is a point $z \in G^\delta(K)$ with $|z - y| < \delta < \epsilon$, hence $\zeta(z) < \gamma(z) + \epsilon \leq \mathrm{dist}(z, \mathrm{sp}(A)) + \epsilon < 2\epsilon < 1$. Thus, M_z is not empty and contains a point which is closer to y than $h(\zeta(z)) + \epsilon \leq h(2\epsilon) + \epsilon \leq u(\epsilon)$. Finally notice that the mapping $(t, \xi) \mapsto h(h(t+\xi) - h(t) + 3\xi) + \xi$ is continuous on the compact set $[0, 1] \times [0, 1]$, hence uniformly continuous. Moreover, for every fixed t it tends to 0 as $\xi \to 0$, thus we can conclude $u(\xi) \to 0$, and we have proved the claim.

Our next goal is the definition of suitable approximating functions for γ and to build the bridge between γ and $\zeta_{m,n}$. Here we have to take into account the following aspects: (i) The functions shall approximate γ locally uniformly. (ii) There shall be a compact set which contains $\mathrm{sp}(A)$ such that all of these functions are greater than 1 outside that set. To do so, define the functions

$$\gamma_m(z) := \min\{\sigma_1((A-zI)P_m|_{\mathrm{Ran}(P_m)}), \sigma_1((A^*-\overline{z}I)P_m|_{\mathrm{Ran}(P_m)})\}$$
$$\gamma_{m,n}(z) := \min\{\sigma_1(P_n(A-zI)P_m|_{\mathrm{Ran}(P_m)}), \sigma_1(P_n(A^*-\overline{z}I)P_m|_{\mathrm{Ran}(P_m)})\}.$$

Now $\sigma_1(P_n(A-zI)P_m|_{\mathrm{Ran}(P_m)}) = \inf\{\|P_n(A-zI)P_m\xi\| : \xi \in \mathrm{Ran}(P_m), \|\xi\| = 1\}$ and $\sigma_1((A-zI)P_m|_{\mathrm{Ran}(P_m)}) = \inf\{\|(A-zI)P_m\xi\| : \xi \in \mathrm{Ran}(P_m), \|\xi\| = 1\}$. Thus, since $P_m \to I$ strongly and $P_{m+1} \geq P_m$, then $\gamma_m \to \gamma$ pointwise and monotonically from above, and by Dini's Theorem the convergence is uniform on every compact set, in particular on the ball $K := B_{m_0}(0)$, with an $m_0 > 2\|A\| + 4$. Also, $\gamma_{m,n} \to \gamma_m$ pointwise monotonically from below as $n \to \infty$, hence again by Dini's Theorem it follows that the convergence is uniform on the ball $K = B_{m_0}(0)$. Outside that ball we have, for $n > m$, by a Neumann argument

$$\begin{aligned}\gamma_{m,n}(z) &= \min\{\sigma_1(P_n(A-zI)P_nP_m), \sigma_1(P_n(A^*-\overline{z}I)P_nP_m)\}\\ &\geq \min\{\sigma_1(P_n(A-zI)P_n), \sigma_1(P_n(A^*-\overline{z}I)P_n)\}\\ &= \|(P_n(A-zI)P_n)^{-1}\|^{-1} = |z|\|(P_n - z^{-1}P_nAP_n)^{-1}\|^{-1} \geq 2.\end{aligned}$$

We can now directly show Step II: For all $n > m \geq m_0$, the points in the finite set $G^{1/m}(B_m(0)) \setminus K$ lead to function values of $\zeta_{m,n}$ being larger than 1, hence $\Gamma_{m,n}(A) = \Upsilon_K^{1/m}(\zeta_{m,n})$. Fix $\epsilon \in (0, 1/2)$. Then there is an $m_1 > m_0$ with $m_1 > 3/\epsilon$ such that $\|\gamma - \gamma_m\|_{\infty,\hat{K}} < \epsilon/3$ on $\hat{K} := B_{h(\mathrm{diam}(K)+2\epsilon)+\epsilon}(0)$ for all $m > m_1$. Moreover, for every m there is an $n_1(m)$ such that $\|\gamma_m - \gamma_{m,n}\|_{\infty,\hat{K}} < \epsilon/3$ for all $n > n_1(m)$. This yields

$$\begin{aligned}\|\gamma - \zeta_{m,n}\|_{\infty,\hat{K}} &\leq \|\gamma - \gamma_m\|_{\infty,\hat{K}} + \|\gamma_m - \gamma_{m,n}\|_{\infty,\hat{K}} + \|\gamma_{m,n} - \zeta_{m,n}\|_{\infty,\hat{K}}\\ &\leq \epsilon/3 + \epsilon/3 + 1/m < \epsilon\end{aligned} \tag{10.7}$$

whenever $m > m_1$ and $n > n_1(m)$. Hence, by the above claim, it holds that $d_{\mathrm{H}}(\Gamma_{m,n}(A), \mathrm{sp}(A)) \leq u(\epsilon)$ whenever $m > m_1$ and $n > n_1(m)$. Since this bound tends to zero as $\epsilon \to 0$, it is proved that

$$\lim_{m\to\infty} \limsup_{n\to\infty} d_{\mathrm{H}}(\Gamma_{m,n}(A), \mathrm{sp}(A)) = 0.$$

To ensure that $(\Gamma_{m,n}(A))_{n\in\mathbb{N}}$ already converges w.r.t. the Hausdorff distance for every fixed m we just mention that the sequence of functions $\{\gamma_{m,n}\}_n$ is monotonic from below, hence $\{\zeta_{m,n}\}_n$ is monotonic as well. Moreover, for every fixed m, these $\zeta_{m,n}$ are effectively evaluated only in finitely many points (namely the points $w \in M_z$ with z in $G^{1/m}(K)$) and can take only finitely many values, where the bounds do not depend on n. Thus the sets $\Gamma_{m,n}(A)$ change only finitely many times as n grows. Consequently, there is an $n_2(m)$ such that all $\Gamma_{m,n}(A)$ with $n \geq n_2(m)$ coincide. This provides the limiting set $\Gamma_m(A) = \lim_{n\to\infty} \Gamma_{m,n}(A)$ and hence we have shown that $\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} \leq 2$.

Step III: $\mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{G}} \geq 2$. To prove that there is no General tower of height one for the self-adjoint case we argue by contradiction and suppose that there is a sequence $\{\Gamma_k\}$ of general algorithms such that $\Gamma_k(A) \to \mathrm{sp}(A)$, and in particular each $\Lambda_{\Gamma_k}(A)$ is finite. Thus, for every A and every k there exists a finite number $N(A,k) \in \mathbb{N}$ such that the evaluations from $\Lambda_{\Gamma_k}(A)$ only take the matrix entries $A_{ij} = \langle Ae_j, e_i\rangle$ with $i, j \leq N(A,k)$ into account. We consider operators of the type

$$A := \bigoplus_{r=1}^{\infty} A_{l_r} \qquad \text{with } \{l_r\} \subset \mathbb{N} \text{ and } A_n := \begin{pmatrix} 1 & & & & 1\\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ 1 & & & & 1\end{pmatrix} \in \mathbb{C}^{n\times n}. \tag{10.8}$$

Then $\mathrm{sp}(A_n) = \{0,2\}$, hence A is bounded, self-adjoint, and $\mathrm{sp}(A) = \{0,2\}$ as well. In order to find a counterexample we simply construct an appropriate sequence $\{l_r\} \subset \mathbb{N}$ by induction: For $C := \mathrm{diag}\{1,0,0,0,\ldots\}$ one obviously has $\mathrm{sp}(C) = \{0,1\}$. Choose $k_0 := 1$ and $l_1 > N(C,k_0)$.

Now, suppose that $l_1,\ldots,l_n$ are already chosen. Then we obviously have that $\mathrm{sp}\,(A_{l_1} \oplus \cdots \oplus A_{l_n} \oplus C) = \{0,1,2\}$, hence there exists a k_n such that

$$\Gamma_k\,(A_{l_1} \oplus \cdots \oplus A_{l_n} \oplus C) \cap B_{\frac{1}{n}}(1) \neq \emptyset$$

for every $k \geq k_n$, where $B_{\frac{1}{n}}(1)$ denotes the closed ball of radius $1/n$ and centre 1. Now, choose

$$l_{n+1} > N(A_{l_1} \oplus \cdots \oplus A_{l_n} \oplus C, k_n) - l_1 - l_2 - \ldots - l_n. \tag{10.9}$$

By this construction, it follows that

$$\Gamma_{k_n}(\oplus_{r=1}^{\infty} A_{l_r}) \cap B_{\frac{1}{n}}(1) = \Gamma_{k_n}(A_{l_1} \oplus \ldots \oplus A_{l_n} \oplus C) \cap B_{\frac{1}{n}}(1) \neq \emptyset \quad \forall\, n \in \mathbb{N}. \tag{10.10}$$

Indeed, since any evaluation function $f_{i,j} \in \Lambda$ just provides the (i,j)-th matrix element, it follows by (10.9) that for any evaluation functions $f_{i,j} \in \Lambda_{\Gamma_{k_n}}(A_{l_1} \oplus \cdots \oplus A_{l_n} \oplus C)$ we have that that

$$f_{i,j}(A_{l_1} \oplus \cdots \oplus A_{l_n} \oplus C) = f_{i,j}(\oplus_{r=1}^{\infty} A_{l_r}).$$

Thus, by assumption (iii) in the definition of a General algorithm (Definition 2.3), we get that $\Lambda_{\Gamma_{k_n}}(A_{l_1} \oplus \cdots \oplus A_{l_n} \oplus C) = \Lambda_{\Gamma_{k_n}}(\oplus_{r=1}^{\infty} A_{l_r})$ which, by assumption (ii) in Definition 2.3, yields (10.10). So, from (10.10), we see that the point 1 is contained in the partial limiting set of the sequence $\{\Gamma_k(\oplus_{r=1}^{\infty} A_{l_r})\}_{k=1}^{\infty}$ which approximates $\mathrm{sp}(A) = \{0,2\}$, a contradiction. Finally, we note that the assertions w.r.t. Ω_i^M are now obvious. $\square$

Remark 10.3. We note that in the case of self-adjoint bounded operators the spectrum $\mathrm{sp}(A)$ is real and the function g can be choosen as $x \mapsto x$. Thus, in the definition of $\Upsilon_K^{\delta}(\zeta)$ it suffices to consider compact $K \subset \mathbb{R}$, the real grid $G^{\delta}(K) := (K + [-\delta,\delta]) \cap (\delta\mathbb{Z})$, and for all $z \in G^{\delta}(K)$ only the two points $z_{1/2} := z \pm \zeta(z)$ in I_z. Also in the case of normal operators, where $g : x \mapsto x$ does the job again, the construction simplifies. In particular, for a given function $\zeta : \mathbb{C} \to [0,\infty)$ we may define sets $\Upsilon_K^{\delta}(\zeta)$ as follows: For $z \in G^{\delta}(K)$ consider $I_z := \{z + \zeta(z)e^{j\delta\mathrm{i}} : j = 0,1,\ldots,\lceil 2\pi\delta^{-1}\rceil\}$ and define $\Upsilon_K^{\delta}(\zeta)$ again as in (10.1). The proof is then the same, up to some obvious adaptions.

Proof of Theorem 3.7. **Step I:** $\mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{G}} \geq 2$. Actually, we will even see that this holds already for the set of all bounded self-adjoint operators. The proof is just an appropriate adaption of the third step in the proof of Theorem 3.3: Assume that there is a sequence $\{\Gamma_k\}$ of general algorithms such that $\Gamma_k(A) \to \mathrm{sp}_{N,\epsilon}(A)$ for all $A \in \Omega_1$, and consider operators of the type (10.8). For sufficiently small ϵ the (N,ϵ)-pseudospectrum is a certain neighbourhood of $\{0,2\}$ disjoint to $B_{\frac{1}{2}}(1)$, independently of the choice of $\{l_r\}$. By exactly the same procedure as before one obtains again that 1 belongs to the partial limiting set of $\Gamma_k(A)$ for a certain A, hence a contradiction.

Step II: $\mathrm{SCI}(\Xi_1,\Omega_2)_{\mathrm{G}} \geq 2$. Recall that Ω_2 denotes the set of bounded operators on $l^2(\mathbb{N})$ whose dispersion is bounded by f. Thus, to show the claim, it suffices to show that for any height one general tower of algorithms $\{\Gamma_n\}_{n\in\mathbb{N}}$ for Ξ_1 there exists a weighted shift S, with $(Su)_1 = 0$ for all $u \in l^{\infty}(\mathbb{N})$ and $Se_n = \alpha_n e_{n+1}$ where $\alpha = \{\alpha_n\}_{n\in\mathbb{N}} \in l^{\infty}(\mathbb{N})$, such that $\Gamma_m(S) \nrightarrow \Xi_1(S)$ when $m \to \infty$, Obviously $S \in \Omega_2$. To construct such an S we let

$$\alpha = \{0,0,\ldots,0,1,0,0,\ldots,0,1,1,0,0,\ldots,0,1,1,1,0,\ldots\}, \qquad \alpha_{l_j+1},\alpha_{l_j+2},\ldots,\alpha_{l_j+j} = 1,$$

for some sequence $\{l_j\}_{j\in\mathbb{N}}$ where $l_{j+1} > l_j + 2j$ that we will determine. Observe that regardless of the choice of $\{l_j\}_{j\in\mathbb{N}}$ we have that $\mathrm{sp}(S) = B_1(0)$ (the closed disc centred at zero with radius one). Indeed, on the one hand $\|S\| = 1$, hence $\mathrm{sp}(S) \subset B_1(0)$. On the other hand, one can define the elementary shift operator $V : e_n \mapsto e_{n+1}$, $n \in \mathbb{N}$, and its left inverse $V^{-} : e_{n+1} \mapsto e_n$, $n \in \mathbb{N}$, $e_1 \mapsto 0$. Then the shifted

copies $(V^-)^{l_j} S V^{l_j}$ converge strongly to the limit operator V whose spectrum $\mathrm{sp}(V) = B_1(0)$ is necessarily contained in the essential spectrum of S (cf. [77] or [63]).

To construct S we will inductively choose $\{l_j\}_{j\in\mathbb{N}}$ with the help of another sequence $\{m_j\}_{j\in\mathbb{Z}_+}$ that will also be chosen inductively. Before we start, define, for any $A \in \Omega_2$ and $m \in \mathbb{N}$, $N(A, m)$ to be the smallest integer so that the evaluations from $\Lambda_{\Gamma_m}(A)$ only take the matrix entries $A_{ij} = \langle Ae_j, e_i\rangle$ with $i, j \le N(A, m)$ into account. Now let $m_0 = 1$ and choose $l_1 > N(0, m_1)$. Suppose that $l_1, \ldots, l_n$ and $m_0, \ldots, m_{n-1}$ are already chosen. Note that $\mathrm{sp}(P_r S) = \{0\}$, since $P_r S = P_r S P_r$ can be regarded as a $r \times r$-triangular matrix with zero-diagonal. Thus, since by assumption $\{\Gamma_m\}_{m\in\mathbb{N}}$ is a General tower of algorithms for Ξ_1, there is an m_n such that $\Gamma_m(P_{l_n+n+1}S) \subset B_{\frac{1}{2}}(0)$, for all $m \ge m_n$. Let

$$l_{n+1} > N(P_{l_n+n+1}S, m_n) \text{ such that also } l_{n+1} > l_n + 2n. \tag{10.11}$$

Then, it follows that $\Gamma_{m_n}(S) = \Gamma_{m_n}(P_{l_{n+1}}S) = \Gamma_{m_n}(P_{l_n+n+1}S)$. Indeed, since any evaluation function $f_{i,j} \in \Lambda$ just provides the (i, j)-th matrix element, it follows by (10.11) that for any evaluation functions $f_{i,j} \in \Lambda_{\Gamma_{m_n}}(S)$ we have that $f_{i,j}(S) = f_{i,j}(P_{l_{n+1}}S) = f_{i,j}(P_{l_n+n+1}S)$. Thus, by assumption (iii) in the definition of a General algorithm (Definition 2.3), we get that $\Lambda_{\Gamma_{m_n}}(S) = \Lambda_{\Gamma_{m_n}}(P_{l_{n+1}}S) = \Lambda_{\Gamma_{m_n}}(P_{l_n+n+1}S)$ which, by assumption (ii) in Definition 2.3 implies the assertion. Thus, by the choice of the sequences $\{l_j\}_{j\in\mathbb{N}}$ and $\{m_j\}_{j\in\mathbb{Z}_+}$, it follows that $\Gamma_{m_n}(S) = \Gamma_{m_n}(P_{l_n+n+1}S) \subset B_{\frac{1}{2}}(0)$ for every n. Since $\mathrm{sp(S)} = B_1(0)$ we observe that $\Gamma_m(S) \nrightarrow \Xi_1(S)$.

Step III: $\mathrm{SCI}(\Xi_1, \Omega_1)_{\mathrm{G}} \ge 3$. This proof requires tools from the section on decision making and logic. The reader is encouraged to read Section 7 before embarking on the proof that can be found in Section 14.

Step IV: $\mathrm{SCI}(\Xi_1, \Omega_1)_{\mathrm{A}} \le 3$ and $\mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{A}} \le 2$. Let $A \in \Omega_1 = \mathcal{B}(l^2(\mathbb{N}))$, and $\epsilon > 0$. We introduce the following continuous functions $\gamma^N : \mathbb{C} \to \mathbb{R}_+$, $\gamma_m^N : \mathbb{C} \to \mathbb{R}_+$ and $\gamma_{m,n}^N : \mathbb{C} \to \mathbb{R}_+$,

$$\gamma^N(z) := \left(\min\left\{\sigma_1\left((A - zI)^{2^N}\right), \sigma_1\left((A^* - \bar{z}I)^{2^N}\right)\right\}\right)^{2^{-N}}$$

$$\gamma_m^N(z) := \left(\min\left\{\sigma_1\left((A - zI)^{2^N} P_m\right), \sigma_1\left((A^* - \bar{z}I)^{2^N} P_m\right)\right\}\right)^{2^{-N}}$$

$$\gamma_{m,n}^N(z) := \left(\min\left\{\sigma_1\left((P_n(A - zI)P_n)^{2^N} P_m\right), \sigma_1\left((P_n(A^* - \bar{z}I)P_n)^{2^N} P_m\right)\right\}\right)^{2^{-N}},$$

where $\sigma_1(B)$ denotes the smallest singular value of B, and in the terms like $\sigma_1(P_n B P_m)$ the operator $P_n B P_m$ is regarded as element of $\mathcal{B}(\mathrm{Ran}(P_m), \mathrm{Ran}(P_n))$. We define the desired approximations $\Gamma_{m,n}(A)$ for $\mathrm{sp}_{N,\epsilon}(A)$ by

$$\Gamma_{m,n}(A) := \{z \in G_m : \gamma_{m,n}^N(z) \le \epsilon\},$$

where $G_m := (m^{-1}(\mathbb{Z} + \mathrm{i}\mathbb{Z})) \cap B_m(0)$. Due to Proposition 10.1 it is clear that the computation of $\Gamma_{m,n}(A)$ requires only finitely many arithmetic operations on finitely many evaluations $\{\langle Ae_j, e_i\rangle : i, j = 1, \ldots, n\}$ of A. Now, one can show that

$$\lim_{m\to\infty} \lim_{n\to\infty} \Gamma_{m,n}(A) = \mathrm{sp}_{N,\epsilon}(A),$$

which proves that $\mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{A}} \le 2$. Furthermore, since $d_{\mathrm{H}}(\mathrm{sp}_{N,\epsilon}(A), \mathrm{sp}(A)) \to 0$ as $\epsilon \to 0$, we arrive at an Arithmetic tower of algorithms of height 3 for Ξ_1 with the algorithms at the lowest level given by

$$\Gamma_{k,m,n}(A) := \{z \in G_m : \gamma_{m,n}^N(z) \le 1/k\}.$$

For the proofs and details we refer to [49].

Step V: $\mathrm{SCI}(\Xi_1, \Omega_2)_{\mathrm{A}} \le 2$ and $\mathrm{SCI}(\Xi_2, \Omega_2)_{\mathrm{A}} = 1$. Let A be such that f is a bound for its dispersion, and $\epsilon > 0$. W.l.o.g. we can assume that $f(n) \ge n$ for every n. Define the composition $g := f \circ \cdots \circ f$ of 2^N copies of f. Besides the already defined functions γ^N, γ_m^N and $\gamma_{m,n}^N$ we additionally introduce $\psi_m^N := \gamma_{m,g(m)}^N$, i.e.

$$\psi_m^N(z) := \left(\min\left\{\sigma_1\left((P_{g(m)}(A - zI)P_{g(m)})^{2^N} P_m\right), \sigma_1\left((P_{g(m)}(A^* - \bar{z}I)P_{g(m)})^{2^N} P_m\right)\right\}\right)^{2^{-N}},$$

and we define the desired approximations $\Gamma_m(A)$ for $\mathrm{sp}_{N,\epsilon}(A)$ by

$$\Gamma_m(A) := \{z \in G_m : \psi_m^N(z) \le \epsilon\}.$$

Again, the computation of $\Gamma_m(A)$ requires only finitely many arithmetic operations on finitely many evaluations $\{\langle Ae_j, e_i\rangle : i, j = 1, \ldots, g(m)\}$ of A.

Obviously, there exists a compact set $K \subset \mathbb{C}$ such that $\gamma_m^N(z) > 2\epsilon$ and $\psi_m^N(z) > 2\epsilon$ for all $z \in \mathbb{C} \setminus K$ and all m. Further note that ψ_m^N converges to γ_m^N uniformly on K. Indeed, since all $z \mapsto (P_{g(m)}(A - zI)P_{g(m)})^{2^N} P_m$ and $z \mapsto (A - zI)^{2^N} P_m$ are operator-valued polynomials of the same degree whose coefficients converge in the norm due to the choice of the function g, we can take into account that $|\sigma_1(B + C) - \sigma_1(B)| \le \|C\|$ holds for arbitrary bounded operators B, C, and we arrive at the conclusion that $|\gamma_m^N(z) - \psi_m^N(z)| \to 0$ as $m \to \infty$ uniformly w.r.t. $z \in K$. In order to simplify the notation we choose a sequence (δ_m) which converges monotonically to zero such that

$$\gamma_m^N(z) + \delta_m \ge \psi_m^N(z) \ge \gamma_m^N(z) - \delta_m \text{ for every } m \text{ and every } z \in K.$$

Moreover, we point out that each of the functions $z \mapsto \psi_m^N(z)$ is continuous on the compact set K, hence even uniformly continuous, and we can assume without loss of generality that, for every m,

$$|\psi_m^N(z) - \psi_m^N(y)| < \delta_m \text{ for arbitrary } z, y \in K,\ |z - y| < 1/m. \tag{10.12}$$

Now let $\Delta_\epsilon(A) := \{z \in \mathbb{C} : \gamma^N(z) \le \epsilon\}$ as well as

$$\Delta_{\epsilon,m}(A) := \{z \in \mathbb{C} : \gamma_m^N(z) \le \epsilon\}, \qquad \Psi_{\epsilon,m}(A) := \{z \in \mathbb{C} : \psi_m^N(z) \le \epsilon\}.$$

By the discussion above, we conclude for all $m \ge k$ that

$$\Delta_{\epsilon+\delta_k,m}(A) \supset \Delta_{\epsilon+\delta_m,m}(A) \supset \Psi_{\epsilon,m}(A) \supset \Delta_{\epsilon-\delta_m,m}(A) \supset \Delta_{\epsilon-\delta_k,m}(A). \tag{10.13}$$

Since, $P_m \le P_{m+1}$ and $P_m \to I$ strongly, $\gamma_m^N \to \gamma^N$ monotonically from above pointwise (and hence locally uniformly by Dini's Theorem). Thus, by [49], $\Delta_{\epsilon+\delta_k,m}(A) \to \Delta_{\epsilon+\delta_k}(A) = \mathrm{sp}_{N,\epsilon+\delta_k}(A)$ and $\Delta_{\epsilon-\delta_k,m}(A) \to \Delta_{\epsilon-\delta_k}(A) = \mathrm{sp}_{N,\epsilon-\delta_k}(A)$ as $m \to \infty$. Hence, since $\mathrm{sp}_{N,\epsilon\pm\delta_k}(A) \to \mathrm{sp}_{N,\epsilon}(A)$ as $k \to \infty$, (10.13) yields $\lim_{m\to\infty} \Psi_{\epsilon,m}(A) = \mathrm{sp}_{N,\epsilon}(A)$. To finish the proof we observe that it is clear that on the one hand $\Psi_{\epsilon,m}(A) \supset \Gamma_m(A)$. On the other hand, for sufficiently large m it holds true that for every point $x \in \Psi_{\epsilon-\delta_m,m}(A)$ there is a point $y_x \in G_m \cap U_{1/m}(x)$ and, by (10.12) we get $|\psi_m^N(y_x) - \psi_m^N(x)| < \delta_m$ that is y_x even belongs to $\Gamma_m(A)$. Thus

$$\Gamma_m(A) + B_{1/m}(0) \supset \Psi_{\epsilon-\delta_m,m}(A)$$

for sufficiently large m. Combining this, we arrive at

$$\Psi_{\epsilon,m}(A) + B_{1/k}(0) \supset \Gamma_m(A) + B_{1/m}(0) \supset \Psi_{\epsilon-\delta_m,m}(A) \supset \Psi_{\epsilon-\delta_k,m}(A),$$

for $m \ge k$ large. By the above, the sets on the left tend to $\mathrm{sp}_{N,\epsilon}(A) + B_{1/k}(0)$ as $m \to \infty$, and the sets on the right converge to $\mathrm{sp}_{N,\epsilon-\delta_k}(A)$ for every k. Since both of these sets converge to $\mathrm{sp}_{N,\epsilon}(A)$ as $k \to \infty$ this provides $\lim_{m\to\infty} \Gamma_m(A) = \mathrm{sp}_{N,\epsilon}(A)$.

The already mentioned fact that $\mathrm{sp}_{N,\epsilon}(A) \to \mathrm{sp}(A)$ as $\epsilon \to 0$ finishes the proof of this Step V.

Step VI: We are left with the compact case. Since $f(n) := n$ is a bound on the dispersion for every compact operator K, we can reuse the algorithms

$$\Gamma_n^{N,\epsilon}(K) := \left\{ z \in G_n : \left(\sigma_1\left((P_n(K - zI)P_n)^{2^N}\right)\right)^{2^{-N}} \le \epsilon \right\}$$

from Step V to obtain an Arithmetic tower of height one for the pseudospectra $\mathrm{sp}_{N,\epsilon}(K)$. Now we claim that

$$\Gamma_n(K) := \Gamma_n^{0,1/n}(K) = \{z \in G_n : \sigma_1\left(P_n(K - zI)P_n\right) \le 1/n\}, \quad n \in \mathbb{N},$$

do the job for the spectrum $\mathrm{sp}(K)$, i.e. for every $\delta > 0$ there exists $n_0 \in \mathbb{N}$ such that $d_{\mathrm{H}}(\mathrm{sp}(K), \Gamma_n(K)) < \delta$ for all $n \ge n_0$. Fix $\delta > 0$ and choose $\epsilon > 0$ such that $\mathrm{sp}_\epsilon(K) \subset \mathrm{sp}(K) + B_{\delta/2}(0)$. Then, due to the above,

there exists $n_1 > 1/\epsilon$ such that $\Gamma_n^{0,1/n}(K) \subset \Gamma_n^{0,\epsilon}(K) \subset \mathrm{sp}_\epsilon(K) + B_{\delta/2}(0) \subset \mathrm{sp}(K) + B_\delta(0)$ for all $n \geq n_1$. Conversely, we know (see e.g. [39, Part II, XI.9 Lemma 5]) that $\mathrm{sp}(P_nKP_n) \to \mathrm{sp}(K)$, i.e. there exists $n_2 \geq n_1$ such that $\mathrm{sp}(K) \subset \mathrm{sp}(P_nKP_n) + B_{\delta/2}(0)$. Choose $n_0 \geq n_2$ such that $n_0 \geq 2\|K\| + 1$ and $n_0 > 2/\delta$. Then, for all $n \geq n_0$ and every $z \in \mathrm{sp}(P_nKP_n)$ there exists $z_n \in G_n$ with $|z - z_n| < 1/n$, hence $\sigma_1(P_n(K - z_nI)P_n) < 1/n$, which implies that $z_n \in \Gamma_n^{0,1/n}(K)$. Thus, $\mathrm{sp}(P_nKP_n) \subset \Gamma_n^{0,1/n}(K) + B_{1/n}(0) \subset \Gamma_n^{0,1/n}(K) + B_{\delta/2}(0)$, and we conclude that $\mathrm{sp}(K) \subset \Gamma_n^{0,1/n}(K) + B_\delta(0)$ for all $n \geq n_0$. $\square$

Proof of Theorem 3.8. **Step I:** $\mathrm{SCI}(\Xi_1, \Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi_1, \Omega_2)_{\mathrm{A}} = 1$. To see that, recall the algorithms (10.3) for the more general setting of Theorem 3.3, and plug in the bound f on the dispersion to obtain the general algorithms $\Gamma_m(A) := \Upsilon^{1/m}_{B_m(0)}(\zeta_m)$, where

$$\zeta_m(z) := \min\{k/m : k \in \mathbb{N},\ k/m \geq \min\{\sigma_1(P_{f(m)}(A - zI)P_m), \sigma_1(P_{f(m)}(A^* - \overline{z}I)P_m)\}\}$$

and $\Upsilon^\delta_K(\zeta)$ is defined in (10.1). Then (10.7) becomes

$$\begin{aligned}\|\gamma - \zeta_m\|_{\infty,\hat{K}} &\leq \|\gamma - \gamma_m\|_{\infty,\hat{K}} + \|\gamma_m - \gamma_{m,f(m)}\|_{\infty,\hat{K}} + \|\gamma_{m,f(m)} - \zeta_m\|_{\infty,\hat{K}} \\ &\leq \epsilon/3 + \epsilon/3 + 1/m < \epsilon\end{aligned}$$

for m larger than a certain $m_2 > m_1$ (cf. Step V in the proof of Theorem 3.3). Again, by the claim in Step II of the proof of Theorem 3.3 we find $d_{\mathrm{H}}(\Gamma_m(A), \mathrm{sp}(A)) < u(\epsilon)$ whenever $m > m_2$. Thus, it is proved that $\lim_{m\to\infty} d_{\mathrm{H}}(\Gamma_m(A), \mathrm{sp}(A)) = 0$ and the rest is obvious.

Step II: $\mathrm{SCI}(\Xi_2, \Omega_2)_{\mathrm{G}} \geq 2$. To see that, let $\{a_n\}_{n\in\mathbb{N}} \subset \{0, 1\}$ be a non-constant sequence and $A = \mathrm{diag}\{a_i\}$ a diagonal operator on $l^2(\mathbb{N})$. Then the spectrum equals $\{0, 1\}$. The essential spectrum contains the point 0 if and only if $\{a_n\}_{n\in\mathbb{N}}$ has infinitely many 0s. Thus, if $\mathrm{SCI}(\Xi_2, \Omega_2)_{\mathrm{G}} = 1$, i.e. there is a General height-one tower of algorithms $\{\Gamma_k\}_{k\in\mathbb{N}}$ which computes the essential spectrum of A, then we can define algorithms for a respective decision problem (e.g. $\tilde{\Xi}(\{a_n\})$: Does $\{a_n\}_{n\in\mathbb{N}}$ contain infinitely many 0s?) via $\tilde{\Gamma}_k(\{a_n\}) =$ *Yes* iff $\mathrm{dist}(0, \Gamma_k(A)) < 1/2$. Thus, $\mathrm{SCI}(\tilde{\Xi})_{\mathrm{G}} = 1$ which contradicts Theorem 7.1.

Step III: $\mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{G}} \geq 3$. This proof requires tools from the section on decision making and logic. The reader is encouraged to read Section 7 before embarking on the proof that can be found in Section 14.

Step IV: $\mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{A}} \leq 3$. We start by defining the following functions on $\mathbb{C}$, where $Q_n := I - P_n$,

$$\begin{aligned}\mu_{m,n,k} &: \lambda \mapsto \min\{\sigma_1(P_k(A - \lambda I)Q_mP_n), \sigma_1(P_k(A - \lambda I)^*Q_mP_n)\} \\ \mu_{m,n} &: \lambda \mapsto \min\{\sigma_1((A - \lambda I)Q_mP_n), \sigma_1((A - \lambda I)^*Q_mP_n)\} \\ \mu_m &: \lambda \mapsto \min\{\sigma_1((A - \lambda I)Q_m), \sigma_1((A - \lambda I)^*Q_m)\}.\end{aligned}$$

Here $P_k(A - \lambda I)Q_mP_n$ is considered as operator on $\mathrm{Ran}(Q_mP_n)$, etc. as usual. Recall from the previous proofs that, for every n, m, $\mu_{m,n,k} \to \mu_{m,n}$ pointwise and monotonically from below as $k \to \infty$ and for every m $\mu_{m,n} \to \mu_m$ pointwise and monotonically from above as $n \to \infty$. Furthermore, $\{\mu_m\}_{m\in\mathbb{N}}$ is pointwise increasing and bounded, hence converges as well. Next, we define the finite grids

$$G_n := \left\{\frac{s + \mathrm{i}t}{2^n} : s, t \in \{-2^{2n}, \dots, 2^{2n}\}\right\},$$

and, for $A \in \Omega_1$,

$$\Gamma_{m,n,k}(A) := \left\{\lambda \in G_n : \mu_{m,n,k}(\lambda) \leq \frac{1}{m}\right\}$$

$$\Gamma_{m,n}(A) := \bigcap_{k\in\mathbb{N}} \Gamma_{m,n,k}(A) = \lim_{k\to\infty} \Gamma_{m,n,k}(A), \tag{10.14}$$

$$\Gamma_m(A) := \overline{\bigcup_{n\in\mathbb{N}} \Gamma_{m,n}(A)} = \lim_{n\to\infty} \Gamma_{m,n}(A), \tag{10.15}$$

$$\Gamma(A) := \bigcap_{m\in\mathbb{N}} \Gamma_m(A) = \lim_{m\to\infty} \Gamma_m(A). \tag{10.16}$$

With Proposition 10.1 it easily follows again that all $\Gamma_{m,n,k}$ are general algorithms in the sense of Definition 2.3 that require only finitely many arithmetic operations. Thus, in order to show that this provides indeed a tower of algorithms for Ξ_2 we need to establish the limits in (10.14), (10.15) and (10.16) and that $\Gamma(A)$ equals $\mathrm{sp}_{\mathrm{ess}}(A)$. To do that we abbreviate $\mathcal{H} := l^2(\mathbb{N})$ and show that

$$\mu(\lambda) := \lim_{m\to\infty} \mu_m(\lambda) \quad \text{equals} \quad \|(A-\lambda I+\mathcal{K}(\mathcal{H}))^{-1}\|^{-1} \quad \text{for all} \quad \lambda \in \mathbb{C}, \tag{10.17}$$

where $A-\lambda I+\mathcal{K}(\mathcal{H})$ denotes the element in the Calkin algebra $\mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ and where we use the convention $\|b^{-1}\|^{-1} := 0$ if the element b is not invertible. Clearly it suffices to consider $\lambda = 0$. The estimate "$\leq$" is trivial in case $\mu(0) = 0$. So, Let $\mu(0) > \epsilon > 0$. Choose $m \in \mathbb{N}$ such that $\mu_m(0) \geq \mu(0) - \epsilon$. The operator $A_0 := AQ_m : \mathrm{Ran}Q_m \to \mathrm{Ran}(AQ_m)$ is invertible, hence the kernel of $A = AQ_m + AP_m$ has finite dimension. $\sigma_1(A^*Q_m) > 0$ yields that $\mathrm{Ran}A$ has finite codimension, hence both A and AQ_m are Fredholm. Let R be the orthogonal projection onto $\mathrm{Ran}AQ_m$, B_0 the inverse of A_0 and $B := B_0R$. Then

$$\begin{aligned} BA - I &= (BA-I)P_m + (BA-I)Q_m = (BA-I)P_m \quad \text{and} \\ AB - I &= (AB-I)(I-R) + (AB-I)R = (AB-I)(I-R) \end{aligned}$$

are compact, i.e. B is a regularizer for A. Now

$$\begin{aligned} \|(A+\mathcal{K}(\mathcal{H}))^{-1}\|^{-1} &\geq \|B\|^{-1} = \|B_0R\|^{-1} \\ &\geq (\|B_0\|\|R\|)^{-1} = \|B_0\|^{-1} = \sigma_1(AQ_m) \geq \mu(0) - \epsilon \end{aligned}$$

gives the estimate "$\leq$" since ϵ is arbitrary.

Conversely, there is nothing to prove if A is not Fredholm, so let $\epsilon > 0$ and $B \in (A+\mathcal{K}(\mathcal{H}))^{-1}$ be a regularizer with $\|B\| \leq \|(A+\mathcal{K}(\mathcal{H}))^{-1}\| + \epsilon$. Since the operator $K := BA - I$ is compact we get for all sufficiently large m that $\|Q_mBAQ_m - Q_m\| = \|Q_mKQ_m\|$ is so small such that $Q_m + Q_mKQ_m$ is invertible in $\mathcal{B}(\mathrm{Ran}(Q_m))$,

$$\underbrace{Q_m(Q_m+Q_mKQ_m)^{-1}Q_mB}_{=:\,B_1\in\mathcal{B}(\mathcal{H})}AQ_m = Q_m \quad \text{and} \quad \|Q_mB - B_1\| < \epsilon.$$

We get that $\sigma_1(AQ_m) > 0$, hence the compression $AQ_m : \mathrm{Ran}(Q_m) \to \mathrm{Ran}(AQ_m)$ is invertible and the compression $B_1|_{\mathrm{Ran}(AQ_m)} : \mathrm{Ran}(AQ_m) \to \mathrm{Ran}(Q_m)$ is its (unique) inverse. Thus, we have $\|B_1\| \geq \|B_1|_{\mathrm{Ran}(AQ_m)}\| = \sigma_1(AQ_m)^{-1}$ and further $\|B\| \geq \|Q_mB\| \geq \|B_1\| - \|Q_mB - B_1\| \geq \sigma_1(AQ_m)^{-1} - \epsilon$. We conclude for sufficiently large m that $\sigma_1(AQ_m)^{-1} \leq \|B\| + \epsilon \leq \|(A+\mathcal{K}(\mathcal{H}))^{-1}\| + 2\epsilon$. Since $\epsilon > 0$ is arbitrary we arrive at $\lim_{m\to\infty}\sigma_1(AQ_m) \geq \|(A+\mathcal{K}(\mathcal{H}))^{-1}\|^{-1}$. Applying this observation to A^* we also find

$$\lim_{m\to\infty} \sigma_1(A^*Q_m) \geq \|(A^*+\mathcal{K}(\mathcal{H}^*))^{-1}\|^{-1} = \|(A+\mathcal{K}(\mathcal{H}))^{-1}\|^{-1},$$

which finishes the proof of (10.17). In particular we now can apply that all of the above functions $\mu_{m,n,k}$, $\mu_{m,n}$, μ_m, μ are continuous w.r.t. λ, and together with the already discussed pointwise monotone convergence results, Dinis Theorem gives that the convergences are even locally uniform.

We can now establish the limits in (10.14), (10.15) and (10.16). Notice that the limits exist since the respective sets under consideration are nested and uniformly bounded: Obviously, $\{\Gamma_{m,n,k}(A)\}_k$ is decreasing. Further $\{\Gamma_{m,n}(A)\}_n$ is increasing since, for every k, $\Gamma_{m,n}(A) \subset \Gamma_{m,n,k}(A) \subset \Gamma_{m,n+1,k}(A)$. Finally, $\{\Gamma_m(A)\}_m$ is decreasing. To see this, choose $z \in \Gamma_m(A)$ and a sequence (z_n) with $z_n \to z$ and $z_n \in \Gamma_{m,n}(A)$, respectively. With the inclusions $\Gamma_{m,n}(A) \subset \Gamma_{m,n,k}(A) \subset \Gamma_{m-1,n,k}(A)$ we conclude $z_n \in \Gamma_{m-1,n}(A)$ for every n, hence $z \in \Gamma_{m-1}(A)$.

We are left with proving that $\Gamma(A)$ equals $\mathrm{sp}_{\mathrm{ess}}(A)$. Let $\lambda \in \mathrm{sp}_{\mathrm{ess}}(A)$. For $m \in \mathbb{N}$, $\mu_m(\lambda) = 0$ and furthermore, there is an $n_0(m)$ and a $\lambda_m \in G_{n_0(m)}$, $|\lambda - \lambda_m| < 1/m$, $\mu_m(\lambda_m) < 1/(2m)$ and $\mu_{m,n}(\lambda_m) < 1/m$ for every $n \geq n_0(m)$. Then, for every k, $\mu_{m,n,k}(\lambda_m) < 1/m$ as well. We conclude that $\lambda_m \in \Gamma_m(A) \subset \Gamma_l(A)$, $l = 1, \ldots, m$. Thus the limit λ of the sequence $\{\lambda_m\}$ belongs to all $\Gamma_l(A)$,

hence to $\Gamma(A)$. Conversely, let $\lambda \notin \mathrm{sp}_{\mathrm{ess}}(A)$. Then $\mu(z) > \epsilon > 0$ for a certain $\epsilon > 0$ and for all z in a certain neighborhood U of λ. Moreover there is an $m_0 > 3/\epsilon$ such that $\mu_m(z) > \epsilon/2$ for all $m \geq m_0$ and $z \in U$, hence $\mu_{m,n}(z) > \epsilon/2$ for all $m \geq m_0$, all n and all $z \in U$. Further, for every $m > m_0$ and n there is a $k_0(m,n)$ such that $\mu_{m,n,k}(z) > \epsilon/3 > 1/m_0 > 1/m$ for all $k \geq k_0(m,n)$ and $z \in U$. Thus, the intersection of U and $\Gamma(A)$ is empty, in particular $\lambda \notin \Gamma(A)$.

Step V: $\mathrm{SCI}(\Xi_2, \Omega_2)_{\mathrm{A}} \leq 2$. Knowing a bound f on the dispersion of A obviously suggests to plug it into the previously defined algorithms, i.e.

$$\hat{\mu}_{m,n} : \lambda \mapsto \min\{\sigma_1(P_{f(n)}(A - \lambda I)Q_m P_n), \sigma_1(P_{f(n)}(A - \lambda I)^* Q_m P_n)\}$$
$$\hat{\Gamma}_{m,n}(A) := \left\{\lambda \in G_n : \hat{\mu}_{m,n}(\lambda) \leq \frac{1}{m}\right\}.$$

Unfortunately, all we know about the functions $\hat{\mu}_{m,n}, \mu_m$ is that they are Lipschitz continuous with Lipschitz constant 1 and that $\hat{\mu}_{m,n}$ converge pointwise to μ_m, but not, whether or when this convergence is monotone. Therefore we have to make a modification in order to guarantee the existence of the desired limiting sets.

Without loss of generality, assume that $f(n) \geq n$ for every n. Let Δ_m denote the square $\Delta_m := \{z \in \mathbb{C} : |\Re(z)|, |\Im(z)| \leq 2^{-(m+1)}\}$ and $\Delta_m(\lambda) := \lambda + \Delta_m$ the respective shifted copies. Moreover, set $Z_m := \{\frac{s+\mathrm{i}t}{2^m} : s,t \in \mathbb{Z}\}$ and

$$\begin{aligned}
S_{m,n}(\lambda) &:= \{i = m+1, \dots, n : \exists z \in \Delta_m(\lambda) \cap G_i : \hat{\mu}_{m,i}(z) \leq 1/m\} \\
T_{m,n}(\lambda) &:= \{i = m+1, \dots, n : \exists z \in \Delta_m(\lambda) \cap G_i : \hat{\mu}_{m,i}(z) \leq 1/(m+1)\} \\
E_{m,n}(\lambda) &:= |S_{m,n}(\lambda)| + |T_{m,n}(\lambda)| - n \\
I_{m,n} &:= \{\lambda \in Z_m : E_{m,n}(\lambda) > 0\} \\
\Gamma_{m,n}(A) &:= \bigcup_{\lambda \in I_{m,n}} \Delta_m(\lambda).
\end{aligned}$$

Roughly speaking, $\Gamma_{m,n}(A)$ is the union of a family of squares $\Delta_m(\lambda)$ with $E_{m,n}(\lambda)$ being positive, which is the case if "most of the $\hat{\mu}_{m,i}$ are small on $\Delta_m(\lambda)$".

To make this precise, we first notice that all $\hat{\mu}_{m,i}(z)$, $i \geq m+1$, with z outside the compact ball $K := B_{2\|A\|+2}(0)$ are larger than one, $I_{m,n}$ are finite, and all $\Gamma_{m,n}(A)$ are contained in K, due to a Neumann argument as in the proof of Theorem 3.3. Further $\hat{\mu}_{m,n} \to \mu_m$ uniformly on K. To see this let $\epsilon > 0$, cover K by $U_{\epsilon/3}(z)$, $z \in K$, and pass to a finite subcovering $U_{\epsilon/3}(z_j)$, $j = 1, \dots, l$. Then choose n large enough such that $|\hat{\mu}_{m,n}(z_j) - \mu_m(z_j)| < \epsilon/3$ for all $j = 1, \dots, l$, and find $|\hat{\mu}_{m,n}(z) - \mu_m(z)| < \epsilon$ for all $z \in K$, due to the Lipschitz continuity of $\hat{\mu}_{m,n}$ and μ_m.

We will show that for each $m \geq 5$ the $E_{m,n}(\lambda)$ are pointwise monotone w.r.t. n for every $\lambda \in Z_m \cap K$, if n is sufficiently large. That is, for every λ there is an $n(\lambda)$ such that either $E_{m,n}(\lambda) \leq 0$ or $E_{m,n}(\lambda) > 0$ for all $n \geq n(\lambda)$. Taking the maximum N of the finite set $\{n(\lambda) : \lambda \in Z_m \cap K\}$ then yields that the $\Gamma_{m,n}(A)$, $n \geq N$, are constant, hence converge as $n \to \infty$. Then we denote the limiting set by $\Gamma_m(A)$, and claim that $\{\Gamma_m(A)\}_m$ is a decreasingly nested sequence, hence converges as well. Indeed, let $z \in \Gamma_{m+1}(A)$, then $z \in \Gamma_{m+1,n}(A)$ for large n, i.e. $z \in \Delta_{m+1}(\lambda)$ for a $\lambda \in I_{m+1,n}$, i.e. $\lambda \in Z_{m+1}$ and $E_{m+1,n}(\lambda) > 0$. Clearly, there exists a $\lambda_0 \in Z_m$ with $\Delta_{m+1}(\lambda) \subset \Delta_m(\lambda_0)$, and further (since $\hat{\mu}_{m,i}(z) \leq \hat{\mu}_{m+1,i}(z)$ holds whenever $n > m+1$)

$$\begin{aligned}
S_{m+1,n}(\lambda) &= \{i = m+2, \dots, n : \exists z \in \Delta_{m+1}(\lambda) \cap G_i : \hat{\mu}_{m+1,i}(z) \leq 1/(m+1)\} \\
&\subset \{i = m+1, \dots, n : \exists z \in \Delta_m(\lambda_0) \cap G_i : \hat{\mu}_{m,i}(z) \leq 1/m\} = S_{m,n}(\lambda_0)
\end{aligned}$$

and analogously $T_{m+1,n}(\lambda) \subset T_{m,n}(\lambda_0)$. Therefore $E_{m+1,n}(\lambda) \leq E_{m,n}(\lambda_0)$, which shows that $\lambda_0 \in I_{m,n}$ and thus $z \in \Gamma_m(A)$.

So, let us now come to the monotonicity of the $\{E_{m,n}(\lambda)\}_n$. For fixed λ and $m \geq 5$ we have to consider three possible cases: The first one is $\mu_m(z) > 1/m$ for all $z \in \Delta_m(\lambda)$. Then there exists an n_0 such

that $\hat{\mu}_{m,n}(z) > 1/m$ for all $n \geq n_0$ and all $z \in \Delta_m(\lambda)$ (take into account that $\Delta_m(\lambda)$ is compact and $\hat{\mu}_{m,n} \to \mu_m$ locally uniformly), hence $|S_{m,n}(\lambda)| + |T_{m,n}(\lambda)|$ is constant and $E_{m,n}(\lambda)$ is monotonically decreasing. Secondly, assume that $\mu_m(z) < 1/m$ for all $z \in \Delta_m(\lambda)$. Then there exists an n_0 such that $\hat{\mu}_{m,n}(z) < 1/m$ for all $n \geq n_0$ and all $z \in \Delta_m(\lambda)$, hence $|S_{m,n}(\lambda)| = n - c$ with a certain constant c, and $E_{m,n}(\lambda) = |T_{m,n}(\lambda)| - c$ is monotonically increasing. Finally, assume that $1/m$ belongs to the interval $[\min\{\mu_m(z) : z \in \Delta_m(\lambda)\}, \max\{\mu_m(z) : z \in \Delta_m(\lambda)\}]$ and notice that the length of that interval is at most 2^{-m} which is less than $1/m - 1/(m+1)$ for $m \geq 5$. Then there exists an n_0 such that $\hat{\mu}_{m,n}(z) > 1/(m+1)$ for all $n \geq n_0$ and all $z \in \Delta_m(\lambda)$, hence $\{|T_{m,n}(\lambda)|\}_{n \geq n_0}$ is constant, and $E_{m,n}(\lambda) = (|S_{m,n}(\lambda)| - n) + |T_{m,n}(\lambda)|$ is monotonically decreasing.

It remains to prove that the final limiting set $\lim_m \Gamma_m(A)$ coincides with the essential spectrum. If $z_0 \in \mathrm{sp}_{\mathrm{ess}}(A)$ then $\mu(z_0) = 0$, hence $\mu_m(z_0) = 0$ for all m. So, for fixed m, we have $\hat{\mu}_{m,n}(z) < 1/(m+1)$ for all sufficiently large n and all z in the neighborhood $U_{1/(2m)}(z_0)$. Choose $\lambda \in Z_m$ such that $z_0 \in \Delta_m(\lambda) \subset U_{1/(2m)}(z_0)$. This is possible since $m \geq 5$. Then it is immediate from the definitions that $E_{m,n}(\lambda) = n - c$ with a constant c for all sufficiently large n, hence $z_0 \in \Gamma_{m,n}(A)$ for m, n large, which yields that $z_0 \in \lim_m \lim_n \Gamma_{m,n}(A)$. Conversely, let $z_0 \notin \mathrm{sp}_{\mathrm{ess}}(A)$, i.e. $\mu(z_0) > 0$. Then, for large m_0, there exists an $\epsilon > 3/m_0$ such that $\mu_m(z_0) > \epsilon$, hence also $\hat{\mu}_{m,n}(z_0) > \epsilon/2$ for $m \geq m_0$ and large n, and then also $\hat{\mu}_{m,n}(z) > \epsilon/3 > 1/m_0$ for all z in a certain neighbourhood U of z_0, which does not dependent on $m \geq m_0$, and all large n. For all sufficiently large $m \geq m_0$ all $\Delta_m(\lambda)$ which contain z_0 are subsets of U, hence $E_{m,n}(\lambda) = d - n$ with a constant d for large n, that is $\Gamma_{m,n}(A) \cap U = \emptyset$. Thus z_0 is not in the limiting set. This finishes the proof. $\square$

11. Proofs of theorems in Section 4

Remark 11.1 (Fourier Transform). In this section we require the Fourier transform on $\mathrm{L}^2(\mathbb{R}^d)$, which will be denoted by $\mathcal{F} : \mathrm{L}^2(\mathbb{R}^d) \to \mathrm{L}^2(\mathbb{R}^d)$. Our definition of $\mathcal{F}$ is as follows:

$$[\mathcal{F}\psi](\xi) = \int_{\mathbb{R}^d} \psi(x) e^{-2\pi i x \cdot \xi}\, dx.$$

For brevity we may write $\hat{\psi}$ instead of $\mathcal{F}\psi$. With this definitions $\mathcal{F}$ is unitary on $\mathrm{L}^2(\mathbb{R}^d)$.

Remark 11.2 (The Attouch-Wets Topology). In (4.2) we introduced the Attouch-Wets metric d_{AW} on the space $\mathcal{M}$ of closed subsets of $\mathbb{C}$. Since it is not convenient to work with d_{AW} directly, we make note of the following simple characterization of convergence w.r.t. d_{AW}. Let $A \subset \mathbb{C}$ and $A_n \subset \mathbb{C}$, $n = 1, 2, \ldots$ be closed and non-empty. Then:

$$d_{\mathrm{AW}}(A_n, A) \to 0 \quad \text{if and only if} \quad d_{\mathcal{K}}(A_n, A) \to 0 \text{ for any compact } \mathcal{K} \subset \mathbb{C}, \tag{11.1}$$

where

$$d_{\mathcal{K}}(S, T) = \max\left\{ \sup_{s \in S \cap \mathcal{K}} d(s, T), \sup_{t \in T \cap \mathcal{K}} d(t, S) \right\}, \tag{11.2}$$

where we use the convention that $\sup_{s \in S \cap \mathcal{K}} d(s, T) = 0$ if $S \cap \mathcal{K} = \emptyset$. We refer to [6, Chapter 3] for details and further discussion. Equivalently, we observe that

$$\begin{aligned} &d_{\mathrm{AW}}(A_n, A) \to 0 \\ &\qquad \text{if and only if} \\ &\forall \delta > 0,\ \mathcal{K} \subset \mathbb{C} \text{ compact}, \exists N \text{ s.t. } \forall n > N,\ A_n \cap \mathcal{K} \subset \mathcal{N}_\delta(A) \text{ and } A \cap \mathcal{K} \subset \mathcal{N}_\delta(A_n) \end{aligned} \tag{11.3}$$

where $\mathcal{N}_\delta(X)$ is the usual open δ-neighborhood of the set X. In this section we will simply use the notation $A_n \to A$ to denote this convergence, since there is no room for confusion.

11.1. The case of bounded potential V.

11.1.1. *Proof of Theorem 4.2.* Before we embark on the proof of the theorem the reader unfamiliar with the concept of Halton sequences may want to review this material. A great reference is [70] (see p. 29 for definition). We will also be needing the following definition and theorem in order to prove Theorem 4.2.

Definition 11.3. *Let $\{t_1, \ldots t_N\}$ be a sequence in $[0,1]^d$. Then we define the* star discrepancy *of $\{t_1, \ldots t_N\}$ to be*

$$D_N^*(\{t_1, \ldots t_N\}) = \sup_{K \in \mathcal{K}} \left| \frac{1}{N} \sum_{k=1}^{N} \chi_K(t_k) - \nu(K) \right|,$$

where $\mathcal{K}$ denotes the family of all subsets of $[0,1]^d$ of the form $\prod_{k=1}^d [0, b_k)$, χ_K denotes the characteristic function on K, $b_k \in (0,1]$ and ν denotes the Lebesgue measure.

Theorem 11.4 ([70]). *If $\{t_k\}_{k\in\mathbb{N}}$ is the Halton sequence in $[0,1]^d$in the pairwise relatively prime bases $b_1, \ldots, b_d$, then*

$$D_N^*(\{t_1, \ldots t_N\}) < \frac{d}{N} + \frac{1}{N} \prod_{k=1}^{d} \left(\frac{b_k - 1}{2\log(b_k)} \log(N) + \frac{b_k + 1}{2} \right) \qquad N \in \mathbb{N}. \tag{11.4}$$

For a proof of this theorem see [70], p. 29. Note that as the right hand side of (11.4) is rather cumbersome to work with, hence it is convenient to define the following constant.

Definition 11.5. *Define $C^*(b_1, \ldots, b_d)$ to be the smallest integer such that for all $N \in \mathbb{N}$*

$$\frac{d}{N} + \frac{1}{N} \prod_{k=1}^{d} \left(\frac{b_k - 1}{2\log(b_k)} \log(N) + \frac{b_k + 1}{2} \right) \leq C^*(b_1, \ldots, b_d) \frac{\log(N)^d}{N}$$

where $b_1, \ldots, b_d$ are as in Theorem 11.4.

Further to these definitions, we shall require a Gabor basis which is the core in the discretisation carried out to produce the tower of algorithms. In particular, let

$$\psi_{k,l}(x) = e^{2\pi i k x} \chi_{[0,1]}(x - l), \qquad k, l \in \mathbb{Z}. \tag{11.5}$$

It is well-known that $\psi_{k,l}$ form an orthonormal basis for $\mathrm{L}^2(\mathbb{R})$. Thus, by applying the Fourier transform,

$$\{\hat{\psi}_{k_1,l_1} \otimes \hat{\psi}_{k_2,l_2} \otimes \cdots \otimes \hat{\psi}_{k_d,l_d} : k_1, l_1, \ldots, k_d, l_d \in \mathbb{Z}\} \tag{11.6}$$

forms an orthonormal basis for $L^2(\mathbb{R}^d)$ since the Fourier transform $\mathcal{F}$ is unitary. Let $\{\varphi_j\}_{j\in\mathbb{N}}$ be an enumeration of the collection of functions above, define

$$\mathcal{S} = \mathrm{span}\{\varphi_j\}_{j\in\mathbb{N}} \tag{11.7}$$

and let

$$\theta : \mathbb{N} \ni j \mapsto (k_1, l_1) \times \ldots \times (k_d, l_d) \in \mathbb{Z}^{2d} \tag{11.8}$$

be the bijection used in this enumeration. Define

$$\begin{aligned} \tilde{k}(m,d) &:= \max\{|k_p| : (k_p, l_p) = \theta(j)_p, p \in \{1, \ldots, d\}, j \in \{1, \ldots, m\}\}, \\ \tilde{l}(m,d) &:= \max\{|l_p| : (k_p, l_p) = \theta(j)_p, p \in \{1, \ldots, d\}, j \in \{1, \ldots, m\}\}, \end{aligned} \tag{11.9}$$

and let

$$C_1(m,d,a) := d^2 \left(4 \frac{(\max\{\tilde{l}(m,d)^2 + \tilde{l}(m,d) + 1/3, 1\})^2}{|a - \tilde{k}(m,d)| + 1} \right)^d, \qquad m, d, a \in \mathbb{N}, \tag{11.10}$$

$$C_2(m,d) := 2^d \left(2((\tilde{l}(m,d) + 1)^4 + \tilde{l}(m,d)^4)^2 (2(\tilde{k}(m,d) + 1) + 2) \right)^d, \qquad m, d \in \mathbb{N}. \tag{11.11}$$

The quantities $C_1(m,d,a)$ and $C_2(m,d)$ may seem to come out of the blue. They stem from Lemma 11.7 and Lemma 11.8 that are technical lemmas needed in order to construct the tower of algorithms. However, $C_1(m,d,a)$ and $C_2(m,d)$ occur in the main proof and thus it is advantageous to introduce them here to prepare the reader.

Remark 11.6 (Assumptions on Λ). As mentioned in Remark 4.1 we will now specify the assumption on the constants in Λ. In particular, Λ will include

$$\{\theta(j)_p : p \leq d, j \in \mathbb{N}\} \cup \{C^*(b_1,\dots,b_d)\} \cup \{\log(k\phi(k))\}_{k=1}^{\infty} \cup \{\phi(k)\}_{k=1}^{\infty},$$

where ϕ is the function describing the bound on the local bounded variation in (4.3). Moreover, Λ will also include

$$\left\{ \frac{\partial^s \hat{\psi}_{k,l}}{\partial \xi^s}(\xi) : \xi \in \mathbb{R},\, k,l \in \mathbb{Z},\, s = 0,2 \right\}. \tag{11.12}$$

Note that it is easy to derive closed form expressions for $\hat{\psi}_{k,l}$ and $\frac{\partial^2 \hat{\psi}_{k,l}}{\partial \xi^2}$, and these expressions will be variations of products of exponential functions and functions of the form $x \mapsto 1/x^p$ for $p = 1,2,3$. For any of the general algorithms $\Gamma : \Omega \to \mathcal{M}$ (where Ω is the appropriate domain), used in the lowest level of the tower, will satisfy assumption II in Remark 2.11. In particular, the constant functions in $\Lambda_\Gamma(A)$ are the same for different inputs $A, B \in \Omega$.

With these preliminaries we are now ready to prove Theorem 4.2 which will be done in several steps.

Proof of $\mathrm{SCI}(\Xi_1,\Omega_2)_{\mathrm{A}} = 1$. We split the proof of this upper bound into several steps to simplify its presentation. In the first two steps we define the tower of algorithms containing Γ_m. In the third step we show that $\lim_{m\to\infty} \Gamma_m = \Xi_1$.

Step I: Defining $\Gamma_m(\{V_\rho\}_{\rho \in \Lambda_{\Gamma_m}(V)})$ and $\Lambda_{\Gamma_m}(V)$. To do that recall $\mathcal{S}$ from (11.7). Note that since $\mathcal{D}(H) = \mathrm{W}^{2,2}(\mathbb{R}^d)$ it is easy to show that $\mathcal{S}$ is a core for H. Let P_m, $m \in \mathbb{N}$, be the projection onto $\mathrm{span}\{\varphi_j\}_{j=1}^m$, and let $z \in \mathbb{C}$. Define

$$S_m(V,z) := (-\Delta + V - zI)P_m \quad \text{and} \quad \tilde{S}_m(V,z) := (-\Delta + \overline{V} - \overline{z}I)P_m.$$

Let

$$\sigma_1(S_m(V,z)) := \min\{(\langle S_m(V,z)f, S_m(V,z)f\rangle)^{\frac{1}{2}} : f \in \mathrm{Ran}(P_m), \|f\| = 1\}$$

and $\sigma_1(\tilde{S}_m(V,z)) := \min\{(\langle \tilde{S}_m(V,z)f, \tilde{S}_m(V,z)f\rangle)^{\frac{1}{2}} : f \in \mathrm{Ran}(P_m), \|f\| = 1\}$, and define

$$\gamma_m(z) := \min\{\sigma_1(S_m(V,z)), \sigma_1(\tilde{S}_m(V,z))\}. \tag{11.13}$$

Note that if we could evaluate γ_m at any point z using only finitely many arithmetic operations of elements of the form $V(x)$, $x \in \mathbb{R}^d$, we could have defined a general algorithm as desired by using $\Upsilon^{1/m}_{B_m(0)}(\gamma_m)$ where $\Upsilon^{1/m}_{B_m(0)}$ is defined in (10.1). Unfortunately, such evaluation is not possible (γ_m may depend on infinitely many samples of V), and we will now focus on finding an approximation to γ_m.

Let $S = \{t_k\}_{k\in\mathbb{N}}$, where $t_k \in [0,1]^d$ is a Halton sequence (see [70] p. 29 for definition) in the pairwise relatively prime bases $b_1,\dots,b_d$ (note that the particular choice of the b_js is not important). Define, for $a > 0$ and $N \in \mathbb{N}$

$$\langle f,g\rangle_{a,N} = \frac{(2a)^d}{N}\sum_{k=1}^{N} f^a(t_k)\overline{g^a(t_k)}, \qquad f,g \in L^2(\mathbb{R}^d)\cap \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)), \tag{11.14}$$

where we have defined the rescaling function on $[0,1]^d$ by

$$f^a = f(a(2\cdot -1),\dots,a(2\cdot -1))|_{[0,1]^d}, \tag{11.15}$$

(we will throughout the proof use the superscript a on a function to indicate (11.15)), where $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)) = \{f : \mathrm{TV}(f|_{[-b,b]^d}) < \infty, \forall b > 0\}$ and $\mathrm{TV}(f|_{[-b,b]^d})$ denotes the total variation, in the sense of Hardy

and Krause (see [70]), of f restricted to $[-b,b]^d$. Note that since $V \in \mathrm{L}^\infty(\mathbb{R}^d) \cap \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)$ and any $f \in \mathrm{Ran}(P_m)$ is smooth we have that $S_m(V,z)f \in L^2(\mathbb{R}^d) \cap \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d))$. Hence, we can define for $n,m \in \mathbb{N}$

$$\begin{aligned}\sigma_{1,n}(S_m(V,z)) &:= \min\{(\langle S_m(V,z)f, S_m(V,z)f\rangle_{n,N(n)})^{\frac{1}{2}} : f \in \mathrm{Ran}(P_m), \|f\| = 1\}\\ \sigma_{1,n}(\tilde{S}_m(V,z)) &:= \min\{(\langle \tilde{S}_m(V,z)f, \tilde{S}_m(V,z)f\rangle_{n,N(n)})^{\frac{1}{2}} : f \in \mathrm{Ran}(P_m), \|f\| = 1\},\end{aligned} \tag{11.16}$$

where $N(n) := \lceil n\phi(n)^4 \rceil$ and where ϕ comes from the definition of Ω_1. Let

$$\zeta_m(z) := \min\{k/m : k \in \mathbb{N},\ k/m \geq \min\{\sigma_{1,n(m)}(S_m(V,z)), \sigma_{1,n(m)}(\tilde{S}_m(V,z))\}\}, \tag{11.17}$$

$$n(m) := \min\{n : \tilde{\tau}(m,n) \leq \frac{1}{m^3}\}, \tag{11.18}$$

and

$$\begin{aligned}\tilde{\tau}(m,n) := &(m+1)mC_1(m,d,n)\\ &+ \big(m^2 + \sigma^2\phi^2(n) + 2(\sigma m + 1)(\phi(n)+1)\big)\\ &\times \big(1 + \sigma^2 + 2\sigma\big)\, C_2(m,d)C^*(b_1,\dots,b_d)\frac{\log(N(n))^d}{N(n)}, \quad N(n) = \lceil n\phi(n)^4 \rceil,\end{aligned} \tag{11.19}$$

$\sigma = 3^d - 2^{d+1} + 2$, $C_1(m,d,n)$ is defined in (11.10), $C_2(m,d)$ is defined in (11.11) and $C^*(b_1,\dots,b_d)$ is defined in Definition 11.5. First, note that the choice of $N(n)$ in (11.19) implies that $\tilde{\tau}(m,n) \to 0$ as $n \to \infty$. Thus, $n(m)$ is well defined. Second, note that it is clear that $\tilde{\tau}$, and hence also $n(m)$, can be evaluated by using finitely many arithmetic operations and comparisons from the set

$$\tilde{\Lambda}_1 = \{\theta(j)_p : p \leq d, j \leq m\} \cup \{C^*(b_1,\dots,b_d)\} \cup \{\log(k\phi(k))\}_{k=1}^r \cup \{\phi(k)\}_{k=1}^r, \tag{11.20}$$

where r is some finite integer depending on m. Recall from Remark 11.6 that we have that $\tilde{\Lambda}_1 \subset \Lambda$.

The function $\tilde{\tau}$ may seem to come somewhat out of the blue, however, it stems from certain bounds in (11.41) (see also (11.42)) on errors of discrete integrals related to (11.16). We can now define

$$\Gamma_m(V) = \Gamma_m(\{V_\rho\}_{\rho\in\Lambda_{\Gamma_m}(V)}) := \Upsilon^{1/m}_{B_m(0)}(\zeta_m),$$

where $\Upsilon^{1/m}_{B_m(0)}(\zeta_m)$ is defined in (10.1) and

$$\Lambda_{\Gamma_m}(V) = \{\rho_x : x \in L_m\} \cup \tilde{\Lambda}_1 \cup \tilde{\Lambda}_2. \tag{11.21}$$

Here $\{\rho_x : x \in L_m\}$ is the set of all point evaluations $\rho_x(V) := V(x)$ at the points in

$$L_m := \{(n(2t_{k,1}-1),\dots,n(2t_{k,d}-1)) : k = 1,\dots,N(n) = \lceil n\phi(n)^4 \rceil, n = n(m)\},$$

where $t_k = \{t_{k,1},\dots,t_{k,d}\}$, $n = n(m)$ is defined in (11.18) and $\tilde{\Lambda}_2$ is a finite set of constant functions that will be determined in (11.27) in Step II.

To show that this provides an arithmetic tower of algorithms for Ξ_1 note that each of the mappings $V \mapsto \Gamma_m(V)$ is an algorithm as desired for arithmetic towers of algorithms. Indeed, $\Lambda_{\Gamma_m}(V)$ is finite and does not depend on V, hence we have a non adaptive tower. Moreover, the computation of $\Upsilon^{1/m}_{B_m(0)}(\zeta_m)$ requires only finitely many evaluations of ζ_m, hence it suffices to demonstrate the following.

Step II: For a single $z \in \mathbb{C}$, the evaluation of $\zeta_m(z)$ requires finitely many arithmetic operations of the elements $\{V_\rho\}_{\rho\in\Lambda_{\Gamma_m}(V)}$. To see this we proceed as follows. For $z \in \mathbb{C}$, form the matrices $Z_m(z), \tilde{Z}_m(z) \in \mathbb{C}^{m\times m}$ by considering the orthonormal basis $\{\varphi_j\}_{j\in\mathbb{N}}$ constructed in the beginning of Step I. More precisely,

$$\begin{aligned}Z_m(z)_{ij} &= \langle S_m(V,z)\varphi_j, S_m(V,z)\varphi_i\rangle_{n,N}, \quad i,j \leq m,\\ \tilde{Z}_m(z)_{ij} &= \langle \tilde{S}_m(V,z)\varphi_j, \widetilde{S}_m(V,z)\varphi_i\rangle_{n,N}, \quad i,j \leq m, \quad N = N(n) = \lceil n\phi(n)^4 \rceil,\end{aligned} \tag{11.22}$$

where $n = n(m)$ is defined in (11.18). Note that forming $Z_m(z)_{ij}$ and $\tilde{Z}_m(z)_{ij}$ require only finitely many arithmetic operations and radicals of the elements $\{V_\rho\}_{\rho\in\Lambda_{\Gamma_m}(V)}$, where we will now specify $\tilde{\Lambda}_2$ in (11.21). Indeed,

$$\begin{aligned}\langle S_m(V,z)\varphi_j, S_m(V,z)\varphi_i\rangle_{n,N} =&\langle \Delta\varphi_j, \Delta\varphi_i\rangle_{n,N} - \langle V\varphi_j, \Delta\varphi_i\rangle_{n,N} - \langle \Delta\varphi_j, V\varphi_i\rangle_{n,N} \\ &+ \langle V\varphi_j, V\varphi_i\rangle_{n,N} - 2\Re(z)(\langle \Delta\varphi_j, \varphi_i\rangle_{n,N} \\ &+ \langle V\varphi_j, \varphi_i\rangle_{n,N}) + |z|^2\langle \varphi_j, \varphi_i\rangle_{n,N}.\end{aligned} \tag{11.23}$$

for $i, j \leq m$. Observe that for $s, t \in \{0, 1\}$, $\tilde{s}, \tilde{t} \in \{0, 2\}$ and $g \in \{V, \overline{V}, |V|^2\}$ it follows that

$$\langle g\,\Delta^s\varphi_j, \Delta^t\varphi_i\rangle_{n,N} = \frac{(2n)^d}{N}\sum_{k=1}^{N}\left(g^n(t_k)\sum_{p\in\Phi(s),q\in\Phi(t)} h^n_{i,j,p,q}(t_k)\right), \qquad i,j\leq m, \tag{11.24}$$

$$\begin{aligned} h_{i,j,p,q}(x) := &\left(\hat{\psi}_{\theta(j)_1}(x_1)\cdots\frac{\partial^{\tilde{s}}\hat{\psi}_{\theta(j)_p}}{\partial x_p^{\tilde{s}}}(x_p)\cdots\hat{\psi}_{\theta(j)_d}(x_d)\right) \\ &\times\overline{\left(\hat{\psi}_{\theta(i)_1}(x_1)\cdots\frac{\partial^{\tilde{t}}\hat{\psi}_{\theta(i)_q}}{\partial x_q^{\tilde{t}}}(x_q)\cdots\hat{\psi}_{\theta(i)_d}(x_d)\right)},\end{aligned} \tag{11.25}$$

$$\Phi(t) = \begin{cases}\{1,\ldots,d\}, & t=1\\ \{1\}, & t=0,\end{cases} \tag{11.26}$$

where $\tilde{s} = 2s$ and $\tilde{t} = 2t$. Note that because of the choice of $\psi_{k,l}$ in (11.5) we have explicit formulas for $\hat{\psi}_{\theta(j)_p}$ and $\frac{\partial^{\tilde{s}}\hat{\psi}_{\theta(j)_p}}{\partial x_p^{\tilde{s}}}$ that are variants of exponential functions. Thus, since $n(m)$ can be evaluated with finitely many arithmetic operations and comparisons of the elements in $\tilde{\Lambda}_1$, and by (11.24), (11.25) and (11.26), it follows that $\langle g\,\Delta^s\varphi_j, \Delta^t\varphi_i\rangle_{n,N}$ can be evaluated by using finitely many arithmetic operations and comparisons of elements in $\{\rho(V) : \rho \in \Lambda_{\Gamma_m}(V)\}$ where $\Lambda_{\Gamma_m}(V)$ is defined in (11.21) and

$$\tilde{\Lambda}_2 = \left\{\rho_x\left(\frac{\partial^{\tilde{s}}\hat{\psi}_{\theta(j)_p}}{\partial x_p^{\tilde{s}}}\right) : x\in L_m,\, 1\leq j\leq m,\, 1\leq p\leq d,\, \tilde{s}=0,2\right\}. \tag{11.27}$$

(As discussed in the assumption in Remark 2.11, we treat the numbers in $\tilde{\Lambda}_2$ as constant functions on Ω). Hence, it follows that forming $Z_m(z)_{ij}$ requires only finitely many arithmetic operations and comparisons of the elements $\{V_\rho\}_{\rho\in\Lambda_{\Gamma_m}(V)}$. The argument for $\tilde{Z}_m(z)_{ij}$ using $\langle \tilde{S}_m(V,z)\varphi_j, \widetilde{S}_m(V,z)\varphi_i\rangle_{a,N}$ is identical.

When $Z_m(z)$ and $\tilde{Z}_m(z)$ are formed, we proceed as follows in order to compute $\zeta_m(z)$. For $k \in \mathbb{N}$, we start with $k = 1$, then:

- Check whether $\min\{\sigma_1(Z_m(z)), \sigma_1(\tilde{Z}_m(z))\} \leq k/m$.
- If not let $k = k + 1$ and repeat, otherwise $\zeta_m(z) = k/m$.

Note that the first step requires finitely many arithmetic operations of $\{Z_m(z)_{ij}\}_{i,j\leq m}$ and $\{\tilde{Z}_m(z)_{ij}\}_{i,j\leq m}$ by Proposition 10.1, and the loop will clearly terminate for a finite k and thus compute $\zeta_m(z)$. Hence, we have proven the assertion that the evaluation of $\zeta_m(z)$ requires finitely many arithmetic operations and comparisons of the elements $\{V_\rho\}_{\rho\in\Lambda_{\Gamma_m}(V)}$ and we conclude that Γ_m are general algorithms as desired for arithmetic towers of algorithms.

Step III: Finally, we show that $\Gamma_m(V) \to \Xi_1(V)$, as $n \to \infty$. Note that, by the properties of the Attouch-Wets topology, and as discussed in Remark 11.2, it suffices to show that for any compact set $\mathcal{K} \subset \mathbb{C}$

$$d_{\mathcal{K}}(\Gamma_m(V), \Xi_1(V)) \longrightarrow 0, \quad n\to\infty, \tag{11.28}$$

where $d_{\mathcal{K}}$ is defined in (11.2). To show (11.28) we start by defining

$$\begin{aligned}\gamma(z) := \min\big\{&\inf\{\|(-\Delta+V-zI)\psi\| : \psi\in \mathrm{W}^{2,2}(\mathbb{R}^d), \|\psi\|=1\},\\ &\inf\{\|(-\Delta+\overline{V}-\overline{z}I)\psi\| : \psi\in \mathrm{W}^{2,2}(\mathbb{R}^d), \|\psi\|=1\}\big\} = \|(-\Delta+V-zI)^{-1}\|^{-1},\end{aligned} \tag{11.29}$$

where we use the convention that $\|(-\Delta+V-zI)^{-1}\|^{-1}=0$ when $z\in\mathrm{sp}(-\Delta+V)$ and proceed similarly to the proof of Theorem 3.3 with the following claim. Before we state the claim recall h from the definition of $\Upsilon^{\delta}_{K}(\zeta)$ in Step I of the proof of Theorem 3.3.

Claim: Let $\mathcal{K}\subset\mathbb{C}$ be any compact set, and let K be a compact set containing $\mathcal{K}$ such that $\mathrm{sp}(-\Delta+V)\cap K\neq\emptyset$ and $0<\delta<\epsilon<1/2$. Suppose that ζ is a function with $\|\zeta-\gamma\|_{\infty,\hat{K}}:=\|(\zeta-\gamma)\chi_{\hat{K}}\|_{\infty}<\epsilon$ on $\hat{K}:=(K+B_{h(\mathrm{diam}(K)+2\epsilon)+\epsilon}(0))$, where $\chi_{\hat{K}}$ denotes the characteristic function of $\hat{K}$. Finally, let u be defined as in (10.4). Then $\lim_{\xi\to0}u(\xi)=0$ and

$$d_{\mathcal{K}}(\Upsilon^{\delta}_{K}(\zeta),\mathrm{sp}(-\Delta+V))\leq u(\epsilon). \tag{11.30}$$

To prove the claim, we first show that

$$\sup_{s\in\Upsilon^{\delta}_{K}(\zeta)\cap\mathcal{K}}\mathrm{dist}(s,\mathrm{sp}(-\Delta+V))\leq u(\epsilon). \tag{11.31}$$

If $\Upsilon^{\delta}_{K}(\zeta)\cap\mathcal{K}=\emptyset$ then there is nothing to prove, thus we assume that $\Upsilon^{\delta}_{K}(\zeta)\cap\mathcal{K}\neq\emptyset$. Let $z\in G^{\delta}(K)$ and recall $G^{\delta}(K)$, h_δ and $I_z=B_{h_\delta(\zeta(z))}(z)\cap(\delta(\mathbb{Z}+\mathrm{i}\mathbb{Z}))$ from the definition of $\Upsilon^{\delta}_{K}(\zeta)$ in Step I of the proof of Theorem 3.3. Notice that we may argue exactly as in (10.5) and deduce that $I_z\subset\hat{K}$. Suppose that $M_z\neq\emptyset$. Note that by (4.4), the monotonicity of h, and the compactness of $\mathrm{sp}(-\Delta+V)\cap K\neq\emptyset$ there is a $y\in\mathrm{sp}(-\Delta+V)$ of minimal distance to z with $|z-y|\leq h(\gamma(z))$. Since $\|\zeta-\gamma\|_{\infty,\hat{K}}<\epsilon$, and by using the monotonicity of h, we get $|z-y|\leq h(\zeta(z)+\epsilon)$. Hence, at least one of the $v\in I_z$, say v_0, satisfies $|v_0-y|<h(\zeta(z)+\epsilon)-h(\zeta(z))+2\delta$. Thus, by noting that $\gamma(v_0)\leq\mathrm{dist}(v_0,-\Delta+V)$, and by the assumption that $\delta<\epsilon$, we get $\zeta(v_0)<\gamma(v_0)+\epsilon<h(\zeta(z)+\epsilon)-h(\zeta(z))+3\epsilon$. By the definition of M_z, this estimate now holds for all points $w\in M_z$. Thus, we may argue exactly as in (10.6) and deduce that $\mathrm{dist}(w,\mathrm{sp}(A))\leq h(h(\zeta(z)+\epsilon)-h(\zeta(z))+3\epsilon)$ which yields (11.31). To see that

$$\sup_{t\in\mathrm{sp}(-\Delta+V)\cap\mathcal{K}}\mathrm{dist}(\Upsilon^{\delta}_{K}(\zeta),t)\leq u(\epsilon), \tag{11.32}$$

(where we assume that $\mathrm{sp}(-\Delta+V)\cap\mathcal{K}\neq\emptyset$) take any $y\in\mathrm{sp}(-\Delta+V)\cap\mathcal{K}\subset K$. Then there is a point $z\in G^{\delta}(K)$ with $|z-y|<\delta<\epsilon$, hence $\zeta(z)<\gamma(z)+\epsilon\leq\mathrm{dist}(z,\mathrm{sp}(-\Delta+V))+\epsilon<2\epsilon<1$. Thus, M_z is not empty and contains a point which is closer to y than $h(\zeta(z))+\epsilon\leq h(2\epsilon)+\epsilon\leq u(\epsilon)$, and this yields (11.32). The fact that $\lim_{\xi\to0}u(\xi)=0$ is shown in Step II of the proof of Theorem 3.3, and we have proved the claim.

Armed with this claim we continue on the path to prove (11.28). We define

$$\gamma_{m,n}(z):=\min\{\sigma_{1,n}(S_m(V,z)),\sigma_{1,n}(\tilde{S}_m(V,z))\}. \tag{11.33}$$

Then $\zeta_m=\gamma_{m,n(m)}$ where $n(m)$ is defined as in (11.18). By Lemma 11.9, $\zeta_m\to\gamma$ locally uniformly, when $m\to\infty$. Let m_0 be large enough so that $B_{m_0}(0)\supset\mathcal{K}$. Then, for all $m\geq m_0$, $\Gamma_m(V)\cap\mathcal{K}=\Upsilon^{1/m}_{B_{m_0}(0)}(\zeta_m)\cap\mathcal{K}$. Choose $K=B_{m_0}(0)$ and $\epsilon\in(0,1/2)$ as in the claim. Then, by the claim, there is an $m_1>m_0$ such that for every $m>m_1$, by (11.30), $d_{\mathcal{K}}(\Gamma_m(V),\Xi_1(V))\leq u(\epsilon)$. Since $\lim_{\xi\to0}u(\xi)=0$ then 11.28) follows. □

To finish the proof of Theorem 4.2 we need to establish the convergence of the functions γ_m, ζ_m and $\gamma_{m,n}$.

Lemma 11.7. *Consider the functions $\gamma_{m,n}$ and γ_m defined in (11.33) and (11.13) respectively. Then $\gamma_{m,n}\to\gamma_m$, locally uniformly as $n\to\infty$.*

Proof. Note that we will be using the notation $\mathrm{TV}_{[-a,a]^d}(f)=\mathrm{TV}(f|_{[-a,a]^d})$. Let, for $s,t\in\{0,1\}$, $i,j\leq m$ and $g\in\{V,\overline{V},|V|^2\}$

$$I(g,\Delta^s\varphi_j,\Delta^t\varphi_i)=\int_{\mathbb{R}^d}g(x)\sum_{p\in\Phi(s),q\in\Phi(t)}h_{i,j,p,q}(x)\,dx,$$

where $h_{i,j,p,q}$ is defined in (11.25) and Φ is defined in (11.26) (recall that $\{\varphi_j\}_{j\in\mathbb{N}}$ is an enumeration of $\{\hat\psi_{k_1,l_1}\otimes\hat\psi_{k_2,l_2}\otimes\cdots\otimes\hat\psi_{k_d,l_d} : k_1,l_1,\ldots,k_d,l_d\in\mathbb{Z}\}$ from (11.6)). Observe that by the definition of $\gamma_{m,n}$ and γ_m in (11.33) and (11.13) the lemma follows if we can show that

$$I(g,\Delta^s\varphi_j,\Delta^t\varphi_i) - \frac{(2n)^d}{N}\sum_{k=1}^{N} g^n(t_k)\sum_{p\in\Phi(s),q\in\Phi(t)} h^n_{i,j,p,q}(t_k)) \longrightarrow 0, \qquad n\to\infty, \tag{11.34}$$

where $N = N(n)$ is from (11.22), $i,j\le m$, $s,t\in\{0,1\}$ and g is either $V,\overline{V},|V|^2$ (recall the notation V^a from (11.15)). Note that, by the multi-dimensional Koksma-Hlawka inequality (Theorem 2.11 in [70]) it follows that

$$\begin{aligned}&|I(g,\Delta^s\varphi_j,\Delta^t\varphi_i) - \frac{(2a)^d}{N}\sum_{k=1}^{N} g^n(t_k)\sum_{p\in\Phi(s),q\in\Phi(t)} h^n_{i,j,p,q}(t_k)|\\ &\le \|g\sum_{p\in\Phi(s),q\in\Phi(t)} h_{i,j,p,q}\chi_{R(n)}\|_{\mathrm{L}^1} + \mathrm{TV}_{[-n,n]^d}\left(g^n\sum_{p\in\Phi(s),q\in\Phi(t)} h^n_{i,j,p,q}\right)D^*_N(t_1,\ldots,t_N),\end{aligned} \tag{11.35}$$

where $R(n) = ([-n,n]^d)^c$. To bound the first part of the right hand side of (11.35) note that

$$\left\|g\sum_{p\in\Phi(s),q\in\Phi(t)} h_{i,j,p,q}\chi_{R(n)}\right\|_{\mathrm{L}^1} \le \|g\|_\infty K_{i,j}(n), \tag{11.36}$$

where

$$K_{i,j}(n) := \sum_{p\in\Phi(s),q\in\Phi(t)}\left\langle |\chi_{([-n,n]^d)^c}\hat\psi_{\theta(j)_1}\cdots\frac{\partial^{\tilde s}\hat\psi_{\theta(j)_p}}{\partial x_p^{\tilde s}}\cdots\hat\psi_{\theta(j)_d}|, |\hat\psi_{\theta(i)_1}\cdots\frac{\partial^{\tilde t}\hat\psi_{\theta(i)_q}}{\partial x_q^{\tilde t}}\cdots\hat\psi_{\theta(i)_d}|\right\rangle,$$

(recall θ from (11.8)) where $\chi_{([-n,n]^d)^c}$ denotes the characteristic function on $([-n,n]^d)^c$. To bound $K_{i,j}(n)$, note that it follows by the definition of $\psi_{k,l}$ with $k,l\in\mathbb{Z}$ in (11.5) and some straightforward integration that for $1\le p\le d$ and $(k_p,l_p)=\theta(j)_p$ we have

$$\left|\hat\psi_{k_p,l_p}(x_p)\right| \le \begin{cases}1 & \text{when } k_p-1\le x_p\le k_p+1,\\ \frac{1}{|x_p-k_p|+1} & \text{otherwise,}\end{cases} \tag{11.37}$$

$$\left|\frac{\partial^2\hat\psi_{k_p,l_p}}{\partial x_p^2}(x_p)\right| \le \begin{cases}l_p^2+l_p+\frac13 & \text{when } k_p-1\le x_p\le k_p+1,\\ \frac{l_p^2+l_p+\frac13}{|x_p-k_p|+1} & \text{otherwise.}\end{cases} \tag{11.38}$$

Hence, if

$$\begin{aligned}\tilde k = \tilde k(m,d) &:= \max\{|k_p| : (k_p,l_p)=\theta(j)_p, p\in\{1,\ldots,d\}, j\in\{1,\ldots,m\}\},\\ \tilde l = \tilde l(m,d) &:= \max\{|l_p| : (k_p,l_p)=\theta(j)_p, p\in\{1,\ldots,d\}, j\in\{1,\ldots,m\}\},\end{aligned}$$

and $n>\tilde k$, then it follows that

$$\begin{aligned}K_{i,j}(n) \le d^2\max\Big\{&\Big\langle|\chi_{([-n,n]^d)^c}\hat\psi_{\theta(j)_1}\cdots\frac{\partial^{2s}\hat\psi_{\theta(j)_p}}{\partial x_p^{2s}}\cdots\hat\psi_{\theta(j)_d}|,\\ &|\hat\psi_{\theta(i)_1}\cdots\frac{\partial^{2t}\hat\psi_{\theta(i)_q}}{\partial x_q^{2t}}\cdots\hat\psi_{\theta(i)_d}|\Big\rangle : p\in\Phi(s), q\in\Phi(t), s,t\in\{0,1\}\Big\}\\ &\le C_1(m,d,n), \qquad C_1(m,d,n) = d^2\left(4\frac{(\max\{\tilde l^2+\tilde l+1/3,1\})^2}{|n-\tilde k|+1}\right)^d.\end{aligned} \tag{11.39}$$

To bound the second part of the right hand side of (11.35) observe that, by Lemma 11.8 we have

$$
\begin{aligned}
&\mathrm{TV}_{[-n,n]^d}\left(g^n \sum_{p\in\Phi(s),q\in\Phi(t)} h^n_{i,j,p,q}\right) \\
&\leq d^2\big(\|g\|_\infty\|h_{i,j,p,q}\|_\infty + \sigma^2\mathrm{TV}_{[-n,n]^d}(g)\mathrm{TV}_{[-n,n]^d}(h_{i,j,p,q}) \\
&\quad + \sigma\left(\mathrm{TV}_{[-n,n]^d}(g)\|h_{i,j,p,q}\|_\infty + \mathrm{TV}_{[-n,n]^d}(h_{i,j,p,q})\|g\|_\infty\right)\big) \\
&\leq \max\left\{\|V\|_\infty, \|V^2\|_\infty, \mathrm{TV}_{[-n,n]^d}(V), \mathrm{TV}_{[-n,n]^d}(|V|^2)\right\}\left(1+\sigma^2+2\sigma\right)C_2(m,d),
\end{aligned}
\tag{11.40}
$$

where $\sigma = 3^d - 2^{d+1} + 2$ and $C_2(m,d)$ is defined in (11.11). Thus, by (11.35), (11.36), (11.39), (11.40), Lemma 11.8 and Theorem 11.4 (recall that $\{t_k\}_{k\in\mathbb{N}}$ is a Halton sequence) we get

$$
\begin{aligned}
&|I(g,\Delta^s\varphi_j,\Delta^t\varphi_i) - \frac{(2n)^d}{N}\sum_{k=1}^{N} V^n(t_k)\sum_{p\in\Phi(s),q\in\Phi(t)} h^n_{i,j,p,q}(t_k)| \\
&\leq \max\{\|V\|_\infty, \|V\|^2_\infty\}C_1(m,d,n) + \max\left\{\|V\|_\infty, \|V^2\|_\infty, \mathrm{TV}_{[-n,n]^d}(V), \mathrm{TV}_{[-n,n]^d}(|V|^2)\right\} \\
&\qquad \times \left(1+\sigma^2+2\sigma\right)C_2(m,d)\left(\frac{d}{N} + \frac{1}{N}\prod_{k=1}^{d}\left(\frac{b_k-1}{2\log(b_k)}\log(N) + \frac{b_k+1}{2}\right)\right) \\
&\leq \tau(\|V\|_\infty, m, n),
\end{aligned}
\tag{11.41}
$$

where the last inequality uses the bound on the total variation of V from (4.3) and

$$
\begin{aligned}
\tau(\|V\|_\infty, m, n) :=& (\|V\|_\infty + 1)\|V\|_\infty C_1(m,d,n) \\
&+ \left(\|V\|^2_\infty + \sigma^2\phi^2(n) + 2(\sigma\|V\|_\infty + 1)(\phi(n)+1)\right) \\
&\times \left(1+\sigma^2+2\sigma\right)C_2(m,d)C^*(b_1,\dots,b_d)\frac{\log(N)^d}{N}, \quad N(n) = \lceil n\phi(n)^4\rceil
\end{aligned}
\tag{11.42}
$$

(recall (11.16)) where $C^*(b_1,\dots,b_d)$ is defined in Definition 11.5. Finally, note that, by the definition of $C_1(m,d,n)$ and the fact that we have chosen $N(n)$ according to (11.42), it follows that $\tau(\|V\|_\infty, m, n) \to 0$ as $n\to\infty$. Hence, (11.34) follows via (11.42), and the proof is finished. $\square$

Lemma 11.8. *For all $a > 0$, $i, j \leq n_2$ and $m, n \leq d$:*

(i) $\mathrm{TV}(h^a_{i,j,m,n}) \leq C_2(m,d)$,
(ii) $\|h^a_{i,j,m,n}\|_\infty \leq C_2(m,d)$,
(iii) *for $g \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^d)$ and $\sigma = 3^d - 2^{d+1} + 2$ we have that*

$$
\begin{aligned}
\mathrm{TV}(g^a h^a_{i,j,p,q}) \leq& \|g\|_\infty\|h_{i,j,p,q}\|_\infty + \sigma^2\mathrm{TV}_{[-a,a]^d}(g)\mathrm{TV}_{[-a,a]^d}(h_{i,j,p,q}) \\
&+ \sigma\left(\mathrm{TV}_{[-a,a]^d}(g)\|h_{i,j,p,q}\|_\infty + \mathrm{TV}_{[-a,a]^d}(h_{i,j,p,q})\|g\|_\infty\right),
\end{aligned}
$$

(iv) $\mathrm{TV}_{[-a,a]^d}(|g|^2) \leq \|g\|^2_\infty + \sigma^2\mathrm{TV}^2_{[-a,a]^d}(g) + 2\sigma\|g\|_\infty\mathrm{TV}_{[-a,a]^d}(g)$

where

$$
C_2(m,d) := 2^d\left(2((\tilde{l}+1)^4 + \tilde{l}^4)^2(2(\tilde{k}+1)+2)\right)^d,
$$

and $\tilde{k}, \tilde{l}$ are defined in (11.9).

Proof. To prove both (i) and (ii) we will use the easy facts that $\mathrm{TV}(h^a_{i,j,p,q}) = \mathrm{TV}_{[-a,a]^d}(h_{i,j,p,q})$ and $\mathrm{TV}(g^a h^a_{i,j,p,q}) = \mathrm{TV}_{[-a,a]^d}(gh_{i,j,p,q})$. To prove (i) of the claim let us first recall (see for example [70], p. 19) that when $\psi \in C^1([-a,a]^d)$ then

$$
\mathrm{TV}_{[-a,a]^d}(\psi) = \sum_{k=1}^{d}\sum_{1\leq i_1<\ldots<i_k\leq d} V^{(k)}(\psi; i_1,\dots,i_k),
\tag{11.43}
$$

where $V^{(k)}(\psi; i_1,\dots,i_k) = V^{(k)}(\psi_{i_1,\dots,i_k})$ and

$$\psi_{i_1,\dots,i_k} : (y_1,\dots,y_k) \mapsto \psi(\tilde{y}_1,\dots,\tilde{y}_d), \qquad \tilde{y}_j = a,\, j \neq i_1,\dots,i_k, \quad \tilde{y}_{i_j} = y_j,$$
$$V^{(k)}(\varphi) = \int_{-a}^{a} \cdots \int_{-a}^{a} \left| \frac{\partial^k \varphi}{\partial x_1 \cdots \partial x_k} \right| \, dx_1 \dots dx_k, \qquad \varphi \in C^1([-a,a]^k).$$

Note that from (11.25) and (11.5) it follows that $h^a_{i,j,p,q} \in C^\infty([0,1]^d)$, so by the definition of h in (11.25) we have that, for $k \in \{1,\dots,d\}$ and $1 \leq i_1 < \dots < i_k \leq d$,

$$\begin{aligned} &V^{(k)}(h^a_{i,j,p,q}; i_1,\dots,i_k) \\ &\leq \prod_{\mu=1}^{d} \max \Bigg[\max_{s,t=0,2} \int_{-a}^{a} \left| \frac{\partial}{\partial x_\mu} \left(\frac{\partial^s \hat{\psi}_{\theta(j)_\mu}}{\partial x_\mu^s}(x_\mu) \overline{\frac{\partial^t \hat{\psi}_{\theta(i)_\mu}}{\partial x_\mu^t}(x_\mu)} \right) \right| dx_\mu, \\ &\qquad \max_{\substack{s,t=0,2 \\ x_\mu \in [-a,a]}} \left| \frac{\partial^s \hat{\psi}_{\theta(j)_\mu}}{\partial x_\mu^s}(x_\mu) \overline{\frac{\partial^t \hat{\psi}_{\theta(i)_\mu}}{\partial x_\mu^t}(x_\mu)} \right| \Bigg], \quad \forall k,p,q \leq d. \end{aligned} \tag{11.44}$$

We will now focus on bounding the right hand side of (11.44). Note that by using the definition of $\psi_{k,l}$ with $k,l \in \mathbb{Z}$ in (11.5) and some straightforward integration it follows that for $1 \leq p \leq d$ and $(k_p,l_p) = \theta(j)_p$ we have

$$\left| \frac{\partial \hat{\psi}_{k_p,l_p}}{\partial x_p}(x_p) \right| \leq \begin{cases} l_p + \frac{1}{2} & \text{when } k_p - 1 \leq x_p \leq k_p + 1, \\ \frac{l+\frac{1}{2}}{|x_p - k_p| + 1} & \text{otherwise,} \end{cases} \tag{11.45}$$

$$\left| \frac{\partial^3 \hat{\psi}_{k_p,l_p}}{\partial x_p^3}(x_p) \right| \leq \begin{cases} \frac{(l_p+1)^4 - l_p^4}{4} & \text{when } k_p - 1 \leq x_p \leq k_p + 1, \\ \frac{(l_p+1)^4 - l_p^4}{4(|x_p - k_p| + 1)} & \text{otherwise.} \end{cases} \tag{11.46}$$

Thus, by using (11.37), (11.38), (11.45) and (11.46) it follows that

$$\begin{aligned} &\max_{s,t=0,2} \int_{-a}^{a} \left| \frac{\partial}{\partial x_\mu} \left(\frac{\partial^s \hat{\psi}_{\theta(j)_\mu}}{\partial x_\mu^s}(x_\mu) \overline{\frac{\partial^t \hat{\psi}_{\theta(i)_\mu}}{\partial x_\mu^t}(x_\mu)} \right) \right| dx_\mu \\ &\qquad \leq 2 \max_{s,t=0,1,2,3} \int_{-\infty}^{\infty} \left| \frac{\partial^s \hat{\psi}_{\theta(j)_\mu}}{\partial x_\mu^s}(x_\mu) \overline{\frac{\partial^t \hat{\psi}_{\theta(i)_\mu}}{\partial x_\mu^t}(x_\mu)} \right| dx_\mu \\ &\qquad \leq 2((\tilde{l}+1)^4 + \tilde{l}^4)^2 \left(2(\tilde{k}+1) + \int_{[-\infty,-1]\cup[1,\infty]} \frac{1}{y^2}\, dy \right) \\ &\qquad = 2((\tilde{l}+1)^4 + \tilde{l}^4)^2 \left(2(\tilde{k}+1) + 2 \right), \end{aligned} \tag{11.47}$$

where $\tilde{k} := \max\{|k_p| : (k_p,l_p) = \theta(j)_p, p \in \{1,\dots,d\}, j \in \{1,\dots,n\}\}$, $\tilde{l} := \max\{|l_p| : (k_p,l_p) = \theta(j)_p, p \in \{1,\dots,d\}, j \in \{1,\dots,n\}\}$. Moreover, by (11.37) and (11.38)

$$\max_{\substack{s,t=0,2 \\ x_\mu \in [-a,a]}} \left| \frac{\partial^s \hat{\psi}_{\theta(j)_\mu}}{\partial x_\mu^s}(x_\mu) \overline{\frac{\partial^t \hat{\psi}_{\theta(i)_\mu}}{\partial x_\mu^t}(x_\mu)} \right| \leq \max\{\tilde{l}^2 + \tilde{l} + 1/3, 1\}, \qquad i,j \leq m, \quad 1 \leq \mu \leq d. \tag{11.48}$$

Hence, from (11.44), (11.47) and (11.48) it follows that for $k \in \{1,\dots,d\}$ and $1 \leq i_1 < \dots < i_k \leq d$,

$$V^{(k)}(h^a_{i,j,p,q}; i_1,\dots,i_k) \leq \left(2((\tilde{l}+1)^4 + \tilde{l}^4)^2 (2(\tilde{k}+1)+2) \right)^d$$

and thus, by (11.43) we get that

$$\begin{aligned} \mathrm{TV}_{[-a,a]^d}(h_{i,j,p,q}) &\leq \left(2((\tilde{l}+1)^4 + \tilde{l}^4)^2 (2(\tilde{k}+1)+2) \right)^d \sum_{k=1}^{d} \binom{d}{k} \\ &\leq 2^d \left(2((\tilde{l}+1)^4 + \tilde{l}^4)^2 (2(\tilde{k}+1)+2) \right)^d, \end{aligned}$$

and thus we have proved (i) in the claim.

To prove (ii) in the claim, we observe that by (11.5), (11.25) and (11.48) it follows that

$$\|h^a_{i,j,p,q}\|_\infty \leq \prod_{\mu=1}^{d} \max_{\substack{s,t=0,2\\ x_\mu\in[-\infty,\infty]}} \left| \frac{\partial^s \hat{\psi}_{\theta(j)_\mu}}{\partial x_\mu^s}(x_\mu) \overline{\frac{\partial^t \hat{\psi}_{\theta(i)_\mu}}{\partial x_\mu^t}(x_\mu)} \right| \leq \left(\max\{\tilde{l}^2 + \tilde{l} + 1/3, 1\}\right)^d,$$

for $i,j \leq m$ and $p,q \leq d$. Obviously, the last part of the above inequality is bounded by $C(m,d)$, which yields the assertion.

To prove (iii) and (iv) we will use the fact (see [13]) that

$$\mathcal{A} = \{f \in \mathcal{M}([-a,a]^d) : \|f\|_\infty + \mathrm{TV}_{[-a,a]^d}(f) < \infty\},$$

where $\mathcal{M}([-a,a]^d)$ denotes the set of measurable functions on $[-a,a]^d$, is a Banach algebra when $\mathcal{A}$ is equipped with the norm $\|f\|_{\mathcal{A}} = \|f\|_\infty + \sigma \mathrm{TV}_{[-a,a]^d}(f)$, where $\sigma > 3^d - 2^{d+1} + 1$. We will let $\sigma = 3^d - 2^{d+1} + 2$. Hence, we get, by the Banach algebra property of the norm and (i) and (ii) that we already have proved, that

$$\begin{aligned} \mathrm{TV}_{[-a,a]^d}(gh_{i,j,p,q}) &\leq \|g\|_\infty \|h_{i,j,p,q}\|_\infty + \sigma^2 \mathrm{TV}_{[-a,a]^d}(g)\mathrm{TV}_{[-a,a]^d}(h_{i,j,p,q}) \\ &\quad + \sigma\left(\mathrm{TV}_{[-a,a]^d}(g)\|h_{i,j,p,q}\|_\infty + \mathrm{TV}_{[-a,a]^d}(h_{i,j,p,q})\|g\|_\infty\right), \quad g \in \mathcal{A}, \end{aligned}$$

finally proving (iii). The proof of (iv) is almost identical. □

Lemma 11.9. *Let ζ_m be defined as in (11.17). Then, $\zeta_m \to \gamma$ locally uniformly, where γ is defined in (11.29).*

Proof. Let γ_m be as defined in (11.13). Also, observe that $\gamma_m \to \gamma$ locally uniformly as $m \to \infty$. Indeed, let $\mathcal{T} = \{\|(-\Delta + V + xI)\psi\| : \psi \in W^{2,2}(\mathbb{R}^d), \|\psi\| = 1\}$. Then, since $\mathcal{S}$ is a core for H (recall $\mathcal{S}$ from Step I of the prof of $\mathrm{SCI}(\Xi_1, \Omega_2)_{\mathrm{A}} = 1$) then every element in $\mathcal{T}$ can be approximated arbitrarily well by $\|(-\Delta + V + xI)\tilde{\varphi}\|$ for some $\tilde{\varphi} \in \mathcal{S}$, thus it follows from (11.29) that we have convergence. Note that the convergence must be monotonically from above by the definition of P_m, and thus Dini's Theorem assures the locally uniform convergence. Thus, it suffices to show that $|\zeta_m - \gamma_m| \to 0$ locally uniformly as $m \to \infty$.

Note that if we define, for $z \in \mathbb{C}$,

$$\begin{aligned} Z_m(z)_{ij} &= \langle S_m(V,z)\varphi_j, S_m(V,z)\varphi_i\rangle_{n,N}, \quad i,j \leq m, \\ \tilde{Z}_m(z)_{ij} &= \langle \tilde{S}_m(V,z)\varphi_j, \widetilde{S}_m(V,z)\varphi_i\rangle_{n,N}, \quad i,j \leq m, \quad N = \lceil n\phi(n)^4 \rceil, \end{aligned}$$

where $n = n(m)$ is defined in (11.18) and

$$W_m(z)_{ij} = \langle S_m(V,z)\varphi_j, S_m(V,z)\varphi_i\rangle, \quad i,j \leq m,$$

$$\tilde{W}_m(z)_{ij} = \langle \tilde{S}_m(V,z)\varphi_j, \widetilde{S}_m(V,z)\varphi_i\rangle, \quad i,j \leq m,$$

the desired convergence follows if we can show that $\|Z_m(z) - W_m(z)\|$ and $\|\tilde{Z}_m(z) - \tilde{W}_m(z)\|$ tend to zero as m tends to infinity for all z in some compact set. However, this follows by the choice of $n(m) = \min\{n : \tilde{\tau}(m,n) \leq \frac{1}{m^3}\}$ in (11.18). In particular, $\tilde{\tau}(m,n) = \tau(m,m,n)$ and clearly $\tau(\|V\|_\infty, m, n) \leq \tau(m,m,n)$ for $\|V\|_\infty \leq m$ (recall τ from (11.42)). Thus it follows immediately by (11.41) that for $z \in K \subset \mathbb{C}$, where K is compact, $\|Z_m(z) - W_m(z)\|_F = \mathcal{O}(\frac{1}{m})$ and $\|\tilde{Z}_m(z) - \tilde{W}_m(z)\|_F = \mathcal{O}(\frac{1}{m})$ for sufficiently large m. Here $\|\cdot\|_F$ denotes the Frobenius norm, and we have shown the desired convergence. □

Proof of $\mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{A}} \leq 2$. The proof stays close to the proof of $\mathrm{SCI}(\Xi_1, \Omega_2)_{\mathrm{A}} = 1$. As before we split the proof into several steps, however, before we embark on the different steps recall that $\Xi_2(V) = \mathrm{sp}_\epsilon(H)$, the setup in Step I of in the proof of $\mathrm{SCI}(\Xi_1, \Omega_2)_{\mathrm{A}} = 1$ and define $\gamma : \mathbb{C} \to [0,\infty)$ as in (11.29), for $m \in \mathbb{N}$, $\gamma_m : \mathbb{C} \to [0,\infty)$ as in (11.13) and, for $m,n \in \mathbb{N}$, $\gamma_{m,n} : \mathbb{C} \to [0,\infty)$ as in (11.33).

Step I: Defining Γ_m, $\Gamma_{m,n}$ and $\Lambda_{\Gamma_{m,n}}(V)$. We start by defining, for $\lambda \in G_m := (m^{-1}(\mathbb{Z}+\mathrm{i}\mathbb{Z}))\cap B_m(0)$,

$$
\begin{aligned}
&S_{m,n}(\lambda) := \{i = m+1,\dots,n : \gamma_{m,i}(\lambda) \le \epsilon - \frac{1}{m}\} \\
&T_{m,n}(\lambda) := \{i = m+1,\dots,n : \gamma_{m,i}(\lambda) \le \frac{1}{(\epsilon - 1/m)^{-1} + 1/m}\} \\
&E_{m,n}(\lambda) := |S_{m,n}(\lambda)| + |T_{m,n}(\lambda)| - n.
\end{aligned} \tag{11.49}
$$

Recall that $\gamma_{m,n}(z) = \min\{\sigma_{1,n}(S_m(V,z)), \sigma_{1,n}(\tilde{S}_m(V,z))\}$, where $\sigma_{1,n}(S_m(V,z))$ and $\sigma_{1,n}(\tilde{S}_m(V,z))$ are defined in (11.16). Thus, we can define

$$
\begin{aligned}
\Gamma_{m,n}(V) = \Gamma_{m,n}(\{V_\rho\}_{\rho\in\Lambda_{\Gamma_{m,n}}(V)}) &:= \{\lambda \in G_m : E_{m,n}(\lambda) > 0\}, \\
\Gamma_m(V) = \lim_{n\to\infty} \Gamma_{m,n}(V)&.
\end{aligned} \tag{11.50}
$$

To determine $\Lambda_{\Gamma_{m,n}}(V)$ we observe that by the definition of ζ_m in (11.17) and the definition of $\gamma_{m,n}$ in (11.33) the evaluation of $\gamma_{m,n}$ requires the same information as the evaluation of ζ_m, with the exception of $\tilde{\Lambda}_1$ in (11.20). Thus,

$$
\Lambda_{\Gamma_{m,n}}(V) = \{\rho_x : \rho_x(V) = V(x), x \in L_n\} \cup \tilde{\Lambda},
$$

where

$$
\tilde{\Lambda} = \left\{ \rho_x \left(\frac{\partial^{\tilde{s}} \hat{\psi}_{\theta(j)_p}}{\partial x_p^{\tilde{s}}} \right) : x \in L_n,\ 1 \le j \le m,\ 1 \le p \le d,\ \tilde{s} = 0, 2 \right\},
$$

$$
L_n := \{(n(2t_{k,1}-1),\dots,n(2t_{k,d}-1)) : k = 1,\dots,\lceil n\phi(n)^4 \rceil\}, \quad t_k = \{t_{k,1},\dots,t_{k,d}\}
$$

and $\psi_{\theta(j)_p}$ is defined in (11.5). Note that showing that the limit in (11.50) exists is part of the proof. Note also that by arguing exactly as in Step II in the proof of $\mathrm{SCI}(\Xi_1,\Omega_2)_{\mathrm{A}} = 1$, it follows that evaluating $\gamma_{m,n}$ requires finitely many arithmetic operations of the elements in $\{V_\rho\}_{\rho\in\Lambda_{\Gamma_{m,n}}(V)}$. Hence, it is clear that Γ_m, $\Gamma_{m,n}$ form an arithmetic tower of algorithms. To finish we only need to show the following.

Step II: We have that

$$
\Gamma_{m,n}(\{V_\rho\}_{\rho\in\Lambda_{\Gamma_{m,n}}(V)}) \longrightarrow \Gamma_m(V), \quad n \to \infty, \tag{11.51}
$$

$$
\Gamma_m(V) \longrightarrow \Xi_2(V), \qquad m \to \infty. \tag{11.52}
$$

We start by showing that $\Gamma_{m,n}(V) = \Gamma_{m,l}(V)$ for all large n and l, thus, the limit exist and we have (11.51). To see this, we start by claiming that for each $\lambda \in G_m$ then either $E_{m,n}(\lambda) \le 0$ for all large n or $E_{m,n}(\lambda) > 0$ for all large n. Observe that if the claim holds this implies that $\Gamma_{m,n}(V) = \Gamma_{m,l}(V)$ for all large n and l. To prove the claim consider the three cases:

$$
(1) : \gamma_m(\lambda) < \frac{1}{(\epsilon - 1/m)^{-1} + 1/m}, \quad (2) : \frac{1}{(\epsilon - 1/m)^{-1} + 1/m} \le \gamma_m(\lambda) \le \epsilon - \frac{1}{m},
$$

$$
(3) : \epsilon - \frac{1}{m} < \gamma_m(\lambda).
$$

In all cases it is the locally uniform convergence $\gamma_{m,n} \to \gamma_m$ established in Lemma 11.7 that is the key. We start with Case 1: In this case $\gamma_{m,n}(\lambda) < \frac{1}{(\epsilon-1/m)^{-1}+1/m}$ for large n, so $|S_{m,n}(\lambda)| = n - c_1$ and $|T_{m,n}(\lambda)| = n - c_2$, $c_1, c_2 \in \mathbb{N}$, yielding the claim. Case 2: This case has two sub-cases. If $\frac{1}{(\epsilon-1/m)^{-1}+1/m} = \gamma_m(\lambda)$ then for large n, $\gamma_{m,n}(\lambda) < \epsilon - \frac{1}{m}$ so $|S_{m,n}(\lambda)| = n - c$ for some positive integer c. Now, either $|T_{m,n}(\lambda)|$ stays constant for large n or it grows. In either case $E_{m,n}(\lambda) = |T_{m,n}(\lambda)| - c$ cannot change sign for large n, thus, the claim holds. When $\frac{1}{(\epsilon-1/m)^{-1}+1/m} < \gamma_m(\lambda) \le \epsilon - \frac{1}{m}$ then $\frac{1}{(\epsilon-1/m)^{-1}+1/m} < \gamma_{m,n}(\lambda)$ for large n, so $|T_{m,n}(\lambda)| = r$, $r \in \mathbb{N}$ for large n. Note that $|S_{m,n}(\lambda)| - n$ cannot change sign for large n, thus, the claim holds. Case 3: In this case $\epsilon - \frac{1}{m} < \gamma_{m,n}(\lambda)$ for large n. Thus, $|S_{m,n}(\lambda)| = c_1$ and $|T_{m,n}(\lambda)| = c_2$

for large n, yielding the claim. As a side note, we observe that due to reasoning in the three cases above we have that

$$\Gamma_m(V) = \{z \in G_m : \gamma_m(z) < \frac{1}{(\epsilon - 1/m)^{-1} + 1/m}\} \cup L, \qquad (11.53)$$
$$L \subset \{z \in G_m : \gamma_m(z) \leq \epsilon - 1/m\}.$$

To see (11.52) we observe that, by definition, $\gamma(z) \leq \gamma_m(z)$ for all $z \in \mathbb{C}$. Thus, we immediately have, by (11.53), that $\Gamma_m(V) \subset \Xi_2(V) = \mathrm{cl}(\{z : \gamma(z) < \epsilon\})$. As we are considering the limit in the Attouch-Wets topology we are only left with the task of showing that for compact set $K \subset \mathbb{C}$ and $\delta > 0$ then $\Xi_2(V) \cap K \subset \mathcal{N}_\delta(\Gamma_m(V))$ for large m. We assume that $\Xi_2(V) \cap K \neq \emptyset$ otherwise there is nothing to prove. To prove the assertion we argue by contradiction. If the statement is untrue, then there is a strictly increasing sequence $\{m_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ such that there exists a sequence $\{z_{m_k}\}_{k\in\mathbb{N}} \subset \Xi_2(V) \cap K$ such that $z_{m_k} \notin \mathcal{N}_\delta(\Gamma_{m_k}(V))$. By compactness of $\Xi_2(V) \cap K$, and by possibly passing to a subsequence, we have that $z_{m_k} \to z$ for some $z \in \Xi_2(V) \cap K$, however, $z \notin \mathcal{N}_{\delta/2}(\Gamma_{m_k}(V))$ for large k. Since $z \in \Xi_2(V) \cap K$ there is a $z_1 \in \mathbb{C}$ such that $\gamma(z_1) < \epsilon$ and $|z - z_1| < \delta/4$. Moreover, by continuity, $\gamma(z) < \epsilon$ on the closed ball $B_\varepsilon(z_1)$ for some $\varepsilon < \delta/4$. In particular, let $s > 0$ be such that $\epsilon - s = \max_{z\in B_\varepsilon(z_1)} \gamma(z)$. Note that by the fact that $\gamma_m \to \gamma$ point wise monotonically from above, and hence locally uniformly (by Dini's theorem), it follows that $\gamma_{m_k}(z) < \epsilon - s_1$ for all large k and $z \in B_\varepsilon(z_1)$ where $0 < s_1 < s$. Note that for large k it follows that $G_{m_k} \cap B_\varepsilon(z_1) \neq \emptyset$ and that $\frac{1}{(\epsilon-1/m_k)^{-1}+1/m_k} > \epsilon - s_1$. Then, for all large k, there exists $\tilde{z}_{m_k} \in G_{m_k} \cap B_\varepsilon(z_1)$ such that $\gamma_{m_k}(\tilde{z}_{m_k}) \leq \frac{1}{(\epsilon-1/m_k)^{-1}+1/m_k}$, thus, $\tilde{z}_{m_k} \in \Gamma_{m_k}(V)$. Also, since $\tilde{z}_{m_k} \in B_\varepsilon(z_1)$ and $|z - z_1| < \delta/4$, it follows that $|\tilde{z}_m - z| < \delta/2$. In particular, $z \in \mathcal{N}_{\delta/2}(\Gamma_{m_k}(V))$ for large k, and we have reached the contradiction. □

Proof of $\mathrm{SCI}(\Xi_1, \Omega_1)_{\mathrm{A}} \leq 2$. The proof is very similar to the proof $\mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{A}} \leq 2$, however, there are important details that differ, so we include the whole proof. We start by defining $\gamma_n = \zeta_n$ where ζ_n is as in Lemma 11.9 and note that $\gamma_n \to \gamma$ locally uniformly.

Step I: Defining Γ_m **and** $\Gamma_{m,n}$. We start by defining, for $\lambda \in G_m := (4^{-m}(\mathbb{Z} + \mathrm{i}\mathbb{Z})) \cap B_m(0)$,

$$S_{m,n}(\lambda) := \{i = m+1, \dots, n : \gamma_i(\lambda) \leq \frac{1}{m}\}$$
$$T_{m,n}(\lambda) := \{i = m+1, \dots, n : \gamma_i(\lambda) \leq \frac{1}{m+1}\}.$$

Now, we can define $E_{m,n}(\lambda)$ as in (11.49) as well as $\Gamma_{m,n}(\{V_\rho\}_{\rho\in\Lambda_{\Gamma_{m,n}}(V)})$ and $\Gamma_m(V)$ as in (11.50). where we will show that the limit in (11.50) exists. Note that, by the definition of γ_n it follows that $\Lambda_{\Gamma_{m,n}}(V)$ is exactly as in (11.21).

Step II: We have that

$$\Gamma_{m,n}(\{V_\rho\}_{\rho\in\Lambda_{\Gamma_{m,n}}(V)}) \longrightarrow \Gamma_m(V), \quad n \to \infty, \qquad (11.54)$$

$$\Gamma_m(V) \longrightarrow \Xi_1(V), \qquad m \to \infty. \qquad (11.55)$$

We start by showing that $\Gamma_{m,n}(V) = \Gamma_{m,l}(V)$ for all large n and l, thus, the limit exist and we have (11.54). To see this, we start by claiming that for each $\lambda \in G_m$ then either $E_{m,n}(\lambda) \leq 0$ for all large n or $E_{m,n}(\lambda) > 0$ for all large n. Observe that if the claim holds this implies that $\Gamma_{m,n}(V) = \Gamma_{m,l}(V)$ for all large n and l. To prove the claim consider the three cases: (1) $\gamma(\lambda) < \frac{1}{m+1}$, (2) $\frac{1}{m+1} \leq \gamma(\lambda) \leq \frac{1}{m}$ and (3) $\frac{1}{m} < \gamma(\lambda)$. In all cases it is the locally uniform convergence $\gamma_n \to \gamma$ that is the key. We start with Case 1: In this case $\gamma_n(\lambda) < \frac{1}{m+1}$ for large n, so $|S_{m,n}(\lambda)| = n - c_1$ and $|T_{m,n}(\lambda)| = n - c_2$, $c_1, c_2 \in \mathbb{N}$, yielding the claim. Case 2: This case has two sub-cases. If $\frac{1}{m+1} = \gamma(\lambda)$ then for large n, $\gamma_n(\lambda) < \frac{1}{m}$ so $|S_{m,n}(\lambda)| = n - c$ for some positive integer c. Now, either $|T_{m,n}(\lambda)|$ stays constant for large n or it grows. In either case $E_{m,n}(\lambda) = |T_{m,n}(\lambda)| - c$ cannot change sign for large n, thus, the claim holds. When $\frac{1}{m+1} < \gamma(\lambda) \leq \frac{1}{m}$ then $\frac{1}{m+1} < \gamma_n(\lambda)$ for large n, so $|T_{m,n}(\lambda)| = r$, $r \in \mathbb{N}$ for large n. Note

that $|S_{m,n}(\lambda)| - n$ cannot change sign for large n, thus, the claim holds. Case 3: In this case $\frac{1}{m} < \gamma_n(\lambda)$ for large n. Thus, $|S_{m,n}(\lambda)| = c_1$ and $|T_{m,n}(\lambda)| = c_2$ for large n, yielding the claim. As a side note, we observe that due to reasoning in the three cases above we have that

$$\Gamma_m(V) = \{z \in G_m : \gamma(z) < \frac{1}{m+1}\} \cup L, \qquad L \subset \{z \in G_m : \gamma(z) \leq \frac{1}{m}\}. \tag{11.56}$$

To show (11.55), let $\delta > 0$ and $K \subset \mathbb{C}$ be compact, we need to show that $\Gamma_m(V) \cap K \subset \mathcal{N}_\delta(\Xi_1(V))$ and $\Xi_1(V) \cap K \subset \mathcal{N}_\delta(\Gamma_m(V))$ for large m. We start with the second inclusion. We assume that $\Xi_1(V) \cap K \neq \emptyset$ otherwise there is nothing to prove. Let m be so large that $4^{-m} < \delta$ and $K \subset B_m(0)$. Let $z \in \Xi_1(V) \cap K$. Then there is a $z_1 \in G_m$ so that $|z_1 - z| \leq 4^{-m} < \delta$. Moreover,

$$z_1 \in \{z : \mathrm{dist}(z, \mathrm{sp}(-\Delta + V)) \leq 4^{-m}\} \subset \{z : \gamma(z) < \frac{1}{m+1}\}.$$

In particular, $z_1 \in \Gamma_m(V)$. Thus, since $|z_1 - z| < \delta$ it follows that $\Xi_1(V) \cap K \subset \mathcal{N}_\delta(\Gamma_m(V))$ as desired. As for the other inclusion, observe that in view of (11.56) and since $\Xi_1(V) = \{z : \gamma(z) = 0\}$ it suffices to show that we may choose an $\epsilon > 0$ such that $\{z : \gamma(z) \leq \epsilon\} \cap K \subset \mathcal{N}_\delta(\{z : \gamma(z) = 0\})$. Suppose not. Then there exists a sequence $\{\epsilon_m\}$, such that $\epsilon_m \to 0$, and a sequence $\{z_m\}$ such that $z_m \in \{z : \gamma(z) \leq \epsilon_m\} \cap K$ but $z_m \notin \mathcal{N}_\delta(\{z : \gamma(z) = 0\})$. By possibly passing to a subsequence we may assume that $z_m \to z_0$ and note that $z_0 \notin \mathcal{N}_\delta(\{z : \gamma(z) = 0\})$. However, by continuity of γ, $\gamma(z_m) \to \gamma(z_0)$, and $\gamma(z_m) \to 0$, hence $\gamma(z_0) = 0$ yielding the contradiction. □

11.2. **The case of unbounded potential V.** In this section we prove Theorem 4.5 on the SCI of spectra and pseudospectra of Schrödinger operators with unbounded potentials. Let us outline the steps of the proof first:

a. *Compactness of the resolvent:* The assumptions on the potential imply that the operator H has a compact resolvent $R(H, z)$ (see Proposition 11.20). Therefore the spectrum is countable consisting of eigenvalues with finite dimensional invariant subspaces.
b. *Finite-dimensional approximations:* The main part of the proof centers around showing that it is possible to construct, with finite amount of evaluations of V, square matrices $\widetilde{H}_n$ whose resolvents (when suitably embedded into the large space) converge to $R(H, z_0)$ in norm at a suitable point z_0 (see Theorem 11.22). Note that this technique is very different from what we used and is only possible due to compactness.
c. *Convergence of the spectrum and pseudospectrum:* We use the convergence at z_0 to show convergence at other points z in the resolvent set.

As the argument is otherwise independent of the particular set-up, we start with a general discussion. In the end we demonstrate the construction of the matrices $\widetilde{H}_n$ and the convergence of the resolvents. We assume the following:

(i) Assumptions on the operator A: Given a closed densely defined operator A in a Hilbert space $\mathcal{H}$ such that at $z_0 \in \mathbb{C}$ the resolvent operator $R(z_0) = (A - z_0)^{-1}$ is compact $R(z_0) \in \mathcal{K}(\mathcal{H})$. Thus $\mathrm{sp}(A) = \{\lambda_j\}$, the spectrum of A, is at most countable with no finite accumulation points.

(ii) Assumptions on the approximations A_n: Suppose A_n is a finite rank approximation to A such that if E_n is the orthogonal projection onto the range of A_n, then $A_n = A_n E_n$. We put further $\mathcal{H}_n = E_n \mathcal{H}$ and denote by $\widetilde{A}_n$ the matrix representing A_n when restricted to the invariant subspace $\mathcal{H}_n$ w.r.t. some orthonormal basis. Now, take the resolvent $(A_n E_n - z E_n)^{-1}$ of this restriction, extend it to $\mathcal{H}_n^\perp$ by zero, and denote this extension by $R_n(z)$. Then $R_n(z) = R_n(z) E_n$, and $R_n(z) = (A_n - z)^{-1} + (I - E_n) z^{-1}$ for all $z \neq 0$ for which the inverse exists. Finally we assume that $R_n(z_0)$ exist and

$$\|R_n(z_0) - R(z_0)\| \longrightarrow 0, \quad n \to \infty. \tag{11.57}$$

Convergence of the spectrum and pseudospectrum. The first step is to conclude that if the finite rank approximations to the resolvent converge in operator norm at one point z_0, then they also converge locally uniformly away from the spectrum of A. To that end denote by $U_r(\mu)$ the open disc at center μ and radius r.

Proposition 11.10. *Suppose $R(z)$ and $R_n(z)$ are as above and satisfy* (11.57)*. Let $\mathcal{K} \subset \mathbb{C}$ be compact, $r > 0$ and define $\mathcal{K}_r = \mathcal{K} \setminus \bigcup_j U_r(\lambda_j)$. Then for large enough n, $R_n(z)$ exists for all $z \in \mathcal{K}_r$ and $\sup_{z\in\mathcal{K}_r} \|R_n(z) - R(z)\| \to 0$ as $n \to \infty$.*

Proof. Clearly $R(z) = R(z_0)(I - (z - z_0)R(z_0))^{-1}$ and $R_n(z) = R_n(z_0)(I - (z - z_0)R_n(z_0))^{-1}$ for all z in which $R(z)$, resp. $R_n(z)$, exist. By (11.57) it suffices to prove the existence of $R_n(z)$ and

$$\sup_{z\in\mathcal{K}_r} \|(I - (z - z_0)R_n(z_0))^{-1} - (I - (z - z_0)R(z_0))^{-1}\| \to 0.$$

However, we know that $(I - (z - z_0)R(z_0))^{-1}$ is meromorphic in the whole plane and hence analytic in the compact set $\mathcal{K}_r$ and in particular uniformly bounded. But this means that it is sufficient to show that the inverses converge, which in turn is immediate from (11.57) since

$$\sup_{z\in\mathcal{K}_r} \|(I - (z - z_0)R_n(z_0)) - (I - (z - z_0)R(z_0))\| = \sup_{z\in\mathcal{K}_r} |z - z_0| \;\; \|R_n(z_0) - R(z_0)\|.$$

To see that this suffices, write $T_n(z) = (I - (z - z_0)R_n(z_0))$, $T(z) = (I - (z - z_0)R(z_0))$ and

$$T_n(z) = T(z)[I + T(z)^{-1}(T_n(z) - T(z))].$$

Then for large enough n and $z \in \mathcal{K}_r$ by a Neumann series argument

$$\|T_n(z)^{-1} - T(z)^{-1}\| \le \|T(z)^{-1}\| \; [(1 - \|T(z)^{-1}\|\|T_n(z) - T(z)\|)^{-1} - 1].$$

□

Proposition 11.11. *Let $\mathcal{K} \subset \mathbb{C}$ be compact and $\delta > 0$. Then, for all large enough n,*

$$\mathrm{sp}(A) \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathrm{sp}(A_n)), \qquad \mathrm{sp}(A_n) \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathrm{sp}(A)).$$

Proof. Since the eigenvalues are exactly the poles of the resolvents, the claim follows immediately from the previous proposition. □

The last proposition gives the convergence of the spectra. The discussion on pseudospectra is somewhat more involved. We need to know that the norm of the resolvent is not constant in any open set. The following is a theorem due to J.Globevnik, E.B.Davies and E.Shargorodsky which we formulate here as a lemma:

Lemma 11.12 ([42] and [31]). *Suppose A is a closed densely defined operator in $\mathcal{H}$ such that the resolvent $R(z) = (A - z)^{-1}$ is compact. Let $\Omega \subset \mathbb{C}$ be open and connected, and assume that, for all $z \in \Omega$, $\|R(z)\| \le M$. Then, for all $z \in \Omega$, $\|R(z)\| < M$. This is particularly true if $\mathcal{H}$ is finite dimensional.*

The theorem in [31] is formulated for Banach spaces X with the extra assumption that X or its dual is complex strictly convex, a condition which holds for Hilbert spaces. The case $\mathcal{H}$ being of finite dimension is already settled by [42]. We put $\gamma(z) = 1/\|R(z)\|$ and $\gamma_n(z) = 1/\|R_n(z)\|$ and summarize the properties of γ and γ_n as follows:

Lemma 11.13. *If (i) and (ii) hold, then $\gamma_n(z) \to \gamma(z)$ uniformly on compact sets. Neither γ, nor γ_n is constant in any open set and they have local minima only where they vanish. Additionally,*

$$\gamma(z) \le \mathrm{dist}(z, \mathrm{sp}(A)). \tag{11.58}$$

Consequently,

$$\mathrm{sp}_\epsilon(A) = \{z \,:\, \gamma(z) \le \epsilon\} = \mathrm{cl}\{z \,:\, \gamma(z) < \epsilon\}, \quad \mathrm{sp}_\epsilon(A_n) = \{z \,:\, \gamma_n(z) \le \epsilon\} = \mathrm{cl}\{z \,:\, \gamma_n(z) < \epsilon\}.$$

Proof. Observe first that (11.58) is just a reformulation of a general property of resolvents. Next, notice that $\|R_n(z)\| = \|R(A_n, z)\|$ and that the norms of resolvents are subharmonic away from spectrum and therefore

γ and γ_n cannot have proper local minima, except when they vanish. Furthermore, they cannot be constant in an open set by Lemma 11.12.

To conclude the local uniform convergence, let M be such that along the curve $\{|z| = M\}$ there are no eigenvalues of A and choose $\mathcal{K}$ as the set $\{|z| \leq M\}$. Choose any ϵ, small enough so that the discs $\{|z - \lambda_j|\} \leq \epsilon/3$ separate the eigenvalues inside $\mathcal{K}$. By Proposition 11.10 we may assume that n is large enough so that for $z \in \mathcal{K}_{\epsilon/3}$ (recall $\mathcal{K}_r$ from Proposition 11.10) we have $|\gamma_n(z) - \gamma(z)| \leq \epsilon/3$. On the other hand, if $|z - \lambda_j| \leq \epsilon/3$ then $\gamma(z) \leq \epsilon/3$ and, since γ_n has to vanish also somewhere in that disc, we have $\gamma_n(z) \leq 2\epsilon/3$ in that disc, hence $|\gamma_n(z) - \gamma(z)| \leq \gamma_n(z) + \gamma(z) \leq \epsilon$. Thus we have $|\gamma_n(z) - \gamma(z)| \leq \epsilon$for all $z \in \mathcal{K}$.

Finally, to justify the equivalence of the characterizations of pseudospectra just notice that the level sets $\{z \,:\, \gamma(z) = \epsilon\}$ and $\{z \,:\, \gamma_n(z) = \epsilon\}$ cannot contain open subsets or isolated points. □

Lemma 11.14. *Assume φ_n and φ are continuous nonnegative functions in $\mathbb{C}$ which have local minima only when they vanish, are not constant in any open set and φ_n converges to φ uniformly in compact sets. Set $\mathcal{S} := \{z \,:\, \varphi(z) \leq 1\}$ and $\mathcal{S}_n := \{z \,:\, \varphi_n(z) \leq 1\}$. Let $\mathcal{K}$ be compact and $\delta > 0$. Then the following hold for all large enough n*

$$\mathcal{S} \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathcal{S}_n), \qquad \mathcal{S}_n \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathcal{S}). \tag{11.59}$$

Proof. Consider the first part of (11.59) and assume that the left hand side is not empty. Due to compactness of $\mathcal{S} \cap \mathcal{K}$ there are points $z_i \in \mathcal{S} \cap \mathcal{K}$ for $i = 1, \ldots, m$ such that $\mathcal{S} \cap \mathcal{K} \subset \bigcup_{i=1}^m U_{\delta/2}(z_i)$. Notice that $\varphi(z_i) \leq 1$. If $\varphi(z_i) < 1$, set $y_i = z_i$. Otherwise, $\varphi(z_i) = 1$, in which case z_i cannot be a local minimum, but since φ is not constant in any open set, there exists a point $y_i \in U_{\delta/2}(z_i)$ such that $\varphi(y_i) < 1$. But since φ_n converges uniformly in compact sets to φ we conclude that for all large enough n and all i we have $\varphi_n(y_i) < 1$. Hence $z_i \in \mathcal{N}_{\delta/2}(\mathcal{S}_n)$ and so $\mathcal{S} \cap \mathcal{K} \subset \bigcup_{i=1}^m U_{\delta/2}(z_i) \subset \mathcal{N}_\delta(\mathcal{S}_n)$.

Consider now the second part of (11.59). If it would not hold, there would exist a sequence $\{n_j\}$ and points $z_{n_j} \in \mathcal{S}_{n_j} \cap \mathcal{K}$ such that $z_{n_j} \notin \mathcal{N}_\delta(\mathcal{S})$. Suppose $z_{n_{j_k}} \to \hat{z}$. Then $\mathrm{dist}(\hat{z}, \mathcal{S}) \geq \delta$ as well. However, writing $\varphi(\hat{z}) \leq |\varphi(\hat{z}) - \varphi(z_{n_{j_k}})| + |\varphi(z_{n_{j_k}}) - \varphi_{n_{j_k}}(z_{n_{j_k}})| + \varphi_{n_{j_k}}(z_{n_{j_k}})$ we obtain $\varphi(\hat{z}) \leq 1$ as the first term on the right tends to zero because φ is continuous, the second term converge to zero as φ_n approximate φ uniformly in compact sets, and $\varphi_{n_{j_k}}(z_{n_{j_k}}) \leq 1$. Hence $\hat{z} \in \mathcal{S} \cap \mathcal{K}$ which is a contradiction. □

Combining these we can state the following result.

Proposition 11.15. *Let $\mathcal{K} \subset \mathbb{C}$ be compact and $\delta > 0$. Then, for all large enough n,*

$$\mathrm{sp}_\epsilon(A) \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathrm{sp}_\epsilon(A_n)), \qquad \mathrm{sp}_\epsilon(A_n) \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathrm{sp}_\epsilon(A)).$$

The general algorithms. Here A, A_n are operators in $\mathcal{H}$ as in (i), (ii) above, while $\widetilde{A}_n$ is the matrix representing A_n when restricted to the finite dimensional invariant subspace $\mathcal{H}_n = E_n\mathcal{H}$. In particular $\|R_n(z)\| = \|(\widetilde{A}_n - z)^{-1}\|$. Denoting by σ_1 the smallest singular value of a square matrix we have $\gamma_n(z) = 1/\|R_n(z)\| = \sigma_1(\widetilde{A}_n - zI)$. Let $r > 0$ and define $G_r := B_r(0) \cap (\frac{1}{2r}(\mathbb{Z} + \mathbf{i}\mathbb{Z}))$. Define Γ_n^1 and Γ_n^2 by

$$\Gamma_n^1(A) = \left\{ z \in G_n : \sigma_1(\widetilde{A}_n - zI) \leq \frac{1}{n} \right\}, \qquad \Gamma_n^2(A) = \left\{ z \in G_n : \sigma_1(\widetilde{A}_n - zI) \leq \epsilon \right\}, \tag{11.60}$$

which we shall prove to be the towers of algorithms for Ξ_1 and Ξ_2 (as defined in Theorem 4.5), respectively. Observe that $\Gamma_n^1(A)$ and $\Gamma_n^2(A)$ can be executed with finite amount of arithmetic operations, if the matrices $\widetilde{A}_n$ are available.

Proposition 11.16. *The algorithms satisfy the following:*

$$\Gamma_n^1(A) \longrightarrow \mathrm{sp}(A), \qquad \Gamma_n^2(A) \longrightarrow \mathrm{sp}_\epsilon(A), \quad n \to \infty. \tag{11.61}$$

Proof. 1. We begin with the second part of (11.61)*.* It suffices to show that given δ and a compact ball $\mathcal{K}$, for large n:

$$\text{(i) } \mathrm{sp}_\epsilon(\widetilde{A}_n) \cap G_n \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathrm{sp}_\epsilon(A)), \qquad \text{(ii) } \mathrm{sp}_\epsilon(A) \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathrm{sp}_\epsilon(\widetilde{A}_n) \cap G_n).$$

The first inclusion follows immediately from Proposition 11.15 . To see (ii) we argue by contradiction and suppose not. Then by possibly passing to an increasing subsequence $\{k_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ there is a sequence $z_n \in (\mathrm{sp}_\epsilon(A) \cap \mathcal{K}) \setminus \mathcal{N}_\delta(\mathrm{sp}_\epsilon(\widetilde{A}_{k_n}) \cap G_{k_n})$ for all n. Since $\mathrm{sp}_\epsilon(A) \cap \mathcal{K}$ is a compact set, by possibly extracting a subsequence, we have that $z_n \to z_0 \in \mathrm{sp}_\epsilon(A) \cap \mathcal{K}$. Consider the open ball $U_{\delta/3}(z_0)$ which must contain all z_n for n sufficiently large. Since $\gamma(z)$ is continuous, positive, not constant in any open set and without nontrivial local minima, it follows that $\mathrm{sp}_\epsilon(A)$ equals the closure of its interior points. In particular $\mathrm{int}(\mathrm{sp}_\epsilon(A)) \cap U_{\delta/3}(z_0) \neq \emptyset$. Suppose then $r > 0$ and y_0 are such that the closure of the open ball $U_r(y_0)$ is inside this open set: $B_r(y_0) \subset \mathrm{int}(\mathrm{sp}_\epsilon(A)) \cap U_{\delta/3}(z_0)$. We claim that $\mathrm{sp}_\epsilon(\widetilde{A}_n) \cap U_r(y_0) = U_r(y_0)$ for all large enough n. Indeed, since $U_r(y_0)$ bounded away from the boundary of the pseudospectrum of A, we have $\gamma(z) \leq \epsilon - s$ for some $s > 0$ and for all $z \in U_r(y_0)$ Now the claim follows from the locally uniform convergence of γ_n.

By the definition of G_n we have that $U_r(y_0) \subset \mathcal{N}_{\delta/3}(U_r(y_0) \cap G_n)$ for large n, so, by the claim, $U_r(y_0) \subset \mathcal{N}_{\delta/3}(\mathrm{sp}_\epsilon(\widetilde{A}_n) \cap G_n)$. Hence, since $U_r(y_0) \subset U_{\delta/3}(z_0)$, it follows that

$$z_n \in U_{\delta/3}(z_0) \subset \mathcal{N}_{2\delta/3}(U_r(y_0)) \subset \mathcal{N}_\delta(\mathrm{sp}_\epsilon(\widetilde{A}_n) \cap G_n),$$

for large n, contradicting $z_n \notin \mathcal{N}_\delta(\mathrm{sp}_\epsilon(\widetilde{A}_n) \cap G_n)$.

2. To prove the first part of (11.61) we argue as follows. Given $\delta > 0$ and compact $\mathcal{K}$, we need to show that for large n:

$$\text{(iii) } \mathrm{sp}(A) \cap \mathcal{K} \subset \mathcal{N}_\delta(\mathrm{sp}_{1/n}(\widetilde{A}_n) \cap G_n) \qquad \text{(iv)) } \mathcal{N}_\delta(\mathrm{sp}(A)) \supset \mathrm{sp}_{1/n}(\widetilde{A}_n) \cap G_n \cap \mathcal{K}.$$

To show (iii), we start by defining $\widetilde{G}_n := \frac{1}{2n}(\mathbb{Z} + \mathbf{i}\mathbb{Z})$ and note that for $\lambda_n \in \mathrm{sp}(\widetilde{A}_n)$ we have that $\mathcal{N}_{1/n}(\{\lambda_n\}) \cap \widetilde{G}_n \neq \emptyset$ for every n. Hence, $\mathrm{sp}(\widetilde{A}_n) \subset \mathcal{N}_{1/n}\left(\mathcal{N}_{1/n}\left(\mathrm{sp}(\widetilde{A}_n)\right) \cap \widetilde{G}_n\right)$. Thus, since $\mathcal{N}_{1/n}(\mathrm{sp}(\widetilde{A}_n)) \subset \mathrm{sp}_{1/n}(\widetilde{A}_n)$, compare (11.58), it follows that $\mathrm{sp}(\widetilde{A}_n) \subset \mathcal{N}_{1/n}\left(\mathrm{sp}_{1/n}(\widetilde{A}_n) \cap \widetilde{G}_n\right)$. Now by the first part of Proposition 11.11 we have that $\mathrm{sp}(A) \cap \mathcal{K} \subset \mathcal{N}_{\delta/2}(\mathrm{sp}(\widetilde{A}_n))$ for large n. Thus, combining the previous observations, we have that

$$\mathrm{sp}(A) \cap \mathcal{K} \subset \mathcal{N}_{\delta/2+1/n}\left(\mathrm{sp}_{1/n}(\widetilde{A}_n) \cap \widetilde{G}_n\right) \subset \mathcal{N}_\delta\left(\mathrm{sp}_{1/n}(\widetilde{A}_n) \cap \widetilde{G}_n\right),$$

for large n. However, since $\mathcal{K}$ is bounded we have that there exists an $r > 0$ such that if $\lambda \in \widetilde{G}_n \cap U_r(0)^c$ then $\mathcal{N}_\delta(\{\lambda\}) \cap \mathrm{sp}(A) \cap \mathcal{K} = \emptyset$ for all n. Hence, $\mathrm{sp}(A) \cap \mathcal{K} \subset \mathcal{N}_\delta\left(\mathrm{sp}_{1/n}(\widetilde{A}_n) \cap G_n\right)$ as desired.

To see (iv), let $r > 0$ be so large that $\mathcal{N}_\delta(U_r(0)^c) \cap \mathcal{K} = \emptyset$. Note that $\mathrm{sp}_\epsilon(A) \to \mathrm{sp}(A)$ as $\epsilon \to 0$. Thus, $\mathrm{sp}_{\epsilon_1}(A) \cap B_r(0) \subset \mathcal{N}_{\delta/2}(\mathrm{sp}(A))$ for a sufficiently small ϵ_1. Also, by the second part of Proposition 11.15 it follows that $\mathrm{sp}_{\epsilon_1}(\widetilde{A}_n) \cap \mathcal{K} \subset \mathcal{N}_{\delta/2}(\mathrm{sp}_{\epsilon_1}(A))$ for large n. However, by the choice of r we have that $\mathrm{sp}_{\epsilon_1}(\widetilde{A}_n) \cap \mathcal{K} \subset \mathcal{N}_{\delta/2}(\mathrm{sp}_{\epsilon_1}(A) \cap B_r(0))$. Clearly, $\mathrm{sp}_{1/n}(\widetilde{A}_n) \cap \mathcal{K} \subset \mathrm{sp}_{\epsilon_1}(\widetilde{A}_n) \cap \mathcal{K}$ for large n. Thus, by patching the above inclusions together we get that

$$\mathrm{sp}_{1/n}(\widetilde{A}_n) \cap \mathcal{K} \subset \mathrm{sp}_{\epsilon_1}(\widetilde{A}_n) \cap \mathcal{K} \subset \mathcal{N}_{\delta/2}(\mathrm{sp}_{\epsilon_1}(A) \cap B_r(0)) \subset \mathcal{N}_\delta(\mathrm{sp}(A)),$$

for large n, as desired. This finishes the proof of Proposition 11.16. □

Next, we pass from these general considerations to the Schrödinger case.

Compactness of the resolvent. We first show that the resolvent of the Schrödinger operator H is compact. To prove this we recall some well known lemmas and definitions from [57].

Definition 11.17. *An operator A on the Hilbert space $\mathcal{H}$ is accretive if the $\mathrm{Re}\langle Ax, x\rangle \geq 0$ for $x \in \mathcal{D}(A)$. It is called m-accretive if there exists no proper accretive extension. If A (possibly after shifting with a scalar)*

is m-accretive and additionally there exists $\beta < \pi/2$ such that $|\arg\langle Ax, x\rangle| \le \beta$ for all $x \in \mathcal{D}(A)$, then A is m-sectorial.

Lemma 11.18 ([57, VI-Theorem 3.3]). *Let A be m-sectorial with $B = \mathrm{Re}\, A$. A has compact resolvent if and only if B has.*

Lemma 11.19 ([57, V-Theorem 3.2]). *If T is closed and the complement of $\mathrm{Num}(T)$ is connected, then for every ζ in the complement of the closure of $\mathrm{Num}(T)$ the following hold: the kernel of $T - \zeta$ is trivial and the range of $T - \zeta$ is closed with constant codimension.*

Proposition 11.20. *Suppose V is continuous $\mathbb{R}^d \to \mathbb{C}$ satisfying the following: $V(x) = |V(x)|e^{i\varphi(x)}$ such that $|V(x)| \to \infty$ as $x \to \infty$, and there exist nonnegative θ_1, θ_2 such that $\theta_1 + \theta_2 < \pi$ and $-\theta_2 \le \varphi(x) \le \theta_1$. Denote by h the operator $h = -\Delta + V$ with domain $\mathcal{D}(h) = C_c^\infty(\mathbb{R}^d)$ and put in $L^2(\mathbb{R}^d)$ $H = h^{**}$. Then $H = -\Delta + V$ is a densely defined operator with a compact resolvent.*

Proof. The proof goes as follows: Notice first that the numerical range of H lies in a sector with opening $2\beta < \pi$. Then we turn the sector into the symmetric position around the positive real axis to get the operator $a(\alpha)$. It is clearly enough to show that $A(\alpha) = a(\alpha)^{**}$ is an m-sectorial operator with half-angle $\beta = (\theta_1 + \theta_2)/2$ which has a compact resolvent. Next, since the numerical range of $a(\alpha)$ is not the whole plane, the operator is closable. Then we conclude that every point away from the numerical range belongs to the resolvent set. This is done based on the fact that the adjoint shares the same key properties as $A(\alpha)$. Then the compactness of the resolvent follows by considering the resolvent of the real part of $A(\alpha)$.

Here is the notation. Put $\alpha = (\theta_1 - \theta_2)/2$ so that $|\alpha| < \pi/2$. Then with

$$\vartheta(x) = \varphi(x) - \alpha \tag{11.62}$$

we have $a(\alpha) := e^{-i\alpha}h = -e^{-i\alpha}\Delta + |V(x)|e^{i\vartheta(x)}$ and after extending $A(\alpha) = a(\alpha)^{**}$, in particular $H(\alpha) := \mathrm{Re}A(\alpha) = -\cos\alpha\, \Delta + \cos\vartheta(x)|V(x)|$.

We claim that the operator $A(\alpha) := e^{-i\alpha}H$ is m-sectorial with half-angle $\beta = (\theta_1 + \theta_2)/2$. Indeed, it is immediate that the numerical range satisfies the following $\mathrm{Num}(a(\alpha)) \subset \{ z = re^{i\theta} : \ |\theta| \le \beta, \ r \ge 0 \}$, which is not the whole complex plane, and we can therefore (by [57, V-Theorem 3.4 on p. 268]) consider the extended closed operator $A(\alpha)$ instead. The next thing is to conclude that points away from this closed sector are in the resolvent set of $A(\alpha)$. Take any point $\zeta = re^{i\varphi}$ with $\beta < |\varphi| \le \pi$, $r > 0$. We need to conclude that $\zeta \notin \mathrm{sp}(A(\alpha))$. Since the complement of $\mathrm{Num}(A(\alpha))$ is connected, the following holds (by Lemma 11.19): the operator $A(\alpha) - \zeta$ has closed range with constant codimension. Thus, we need that the range is the whole space. Put for that purpose $T = A(\alpha) - \zeta$. Suppose there is $g \ne 0$ such that $g \in \mathrm{Ran}(T)^\perp$. Then for all $f \in \mathcal{D}(T)$ we have $\langle Tf, g\rangle = 0$ which means, as $\mathcal{D}(T)$ is dense, that $T^*g = 0$. But that is not the case as $A(\alpha)^* - \overline{\zeta}$ is also closed whose complement of the numerical range is connected and hence does not have a nontrivial kernel.

The proof of Proposition 11.20 can now be completed by invoking Lemma 11.18 since it is well known ([78], Theorem XIII.67) that (since $\alpha < \pi/2$) the self-adjoint operator $H(\alpha)$ has compact resolvent when the potential $|V(x)|$ tends to infinity with x. $\square$

We shall next consider the discretisation of H and of $A(\alpha)$. It shall be clear that the discrete versions have their numerical ranges inside the same sectors, where the numerical range of an operator T is denoted by $\mathrm{Num}(T)$. Thus all resolvents can be estimated using the fact that if $(T - \zeta)^{-1}$ is regular outside the closure of $\mathrm{Num}(T)$, then $\|(T - \zeta)^{-1}\| \le 1/\mathrm{dist}(\zeta, \mathrm{Num}(T))$.

Discretizing the Schrödinger operator. We shall show how to assemble the matrices $\widetilde{H}_n$ mentioned above. The underlying Hilbert space is again $L^2(\mathbb{R}^d)$ and we start with approximating the Laplacian. Let

$1 \le j \le d,\ t \in \mathbb{R}$ and define $U_{j,t}$ to be the one-parameter unitary group of translations

$$U_{j,t}\psi(x_1,\dots,x_d) = \psi(x_1,\dots,x_j - t,\dots,x_d)$$

and let P_j be the infinitesimal generator of $U_{j,t}$ so that $U_{j,t} = e^{\mathrm{i}tP_j}$ and $P_j = \lim_{t\to 0}\frac{1}{\mathrm{i}t}(U_{j,t} - I)$. Thus, defining $\Phi_n(x) = \frac{n}{\mathrm{i}}(e^{\mathrm{i}\frac{1}{n}x} - 1)$ with $n \in \mathbb{N}$ and $x \in \mathbb{R}$, it follows that

$$|\Phi_n|^2(P_j)\psi(x) = n^2(-\psi(x_1,\dots,x_j + 1/n,\dots x_d) - \psi(x_1,\dots,x_j - 1/n,\dots x_d) + 2\psi(x)) \tag{11.63}$$

is the discretized Laplacian in the j direction. The full discretized Laplacian is therefore $\sum_{j=1}^d |\Phi_n|^2(P_j)$. Now we replace V by an appropriate approximation. Consider the lattice $(\frac{1}{n}\mathbb{Z})^d$ as a subset of $\mathbb{R}^d$ and for $y \in (\frac{1}{n}\mathbb{Z})^d$ define the box

$$Q_n(y) = \left\{ x = (x_1,\dots,x_d) : x_j \in \left[y_j - \frac{1}{2n}, y_j + \frac{1}{2n}\right),\ 1 \le j \le d \right\}. \tag{11.64}$$

Let $S_n = [-\lfloor\sqrt{n}\rfloor, \lfloor\sqrt{n}\rfloor]^d \subset \mathbb{R}^d$ and define E_n to be the orthogonal projection onto the subspace

$$\left\{ \psi \in L^2(\mathbb{R}^d) : \psi = \sum_{y\in(\frac{1}{n}\mathbb{Z})^d\cap S_n} \alpha_y \chi_{Q_n(y)},\ \alpha_y \in \mathbb{C} \right\}, \tag{11.65}$$

where $\chi_{Q_n(y)}$ denotes the characteristic function on $Q_n(y)$. Define the approximate potential as

$$V_n(x) = \begin{cases} V(y) & x \in Q_n(y) \cap S_n \text{ for some } y \in (\frac{1}{n}\mathbb{Z})^d, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $V_n = E_n V_n E_n$, but that, in general, $V_n \neq E_n V E_n$. Finally, we define the approximate Schrödinger operator $H_n : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ defined as

$$H_n = E_n \sum_{j=1}^d |\Phi_n|^2(P_j) E_n + V_n. \tag{11.66}$$

Remark 11.21. Note that the restriction $H_n|_{\mathrm{Ran}(E_n)}$ of H_n to the image of E_n has a matrix representation $\widetilde{H}_n \in \mathbb{C}^{m\times m}$ (where $m = \dim(\mathrm{Ran}(E_n))$) defined as follows. First, for $y_1, y_2 \in (\frac{1}{n}\mathbb{Z})^d \cap S_n$,

$$\langle |\Phi_n|^2(P_j)E_n n^{d/2}\chi_{Q_n(y_1)}, n^{d/2}\chi_{Q_n(y_2)}\rangle = \begin{cases} 2n^2 & y_1 = y_2 \\ -n^2 & y_1 - y_2 = \pm 1/n e_j \\ 0 & \text{otherwise} \end{cases}$$

and $\langle V_n n^{d/2}\chi_{Q_n(y_1)}, n^{d/2}\chi_{Q_n(y_2)}\rangle = V(y_1)$ when $y_1 = y_2$ and zero otherwise. Thus, we can form the matrix representation of $H_n|_{\mathrm{Ran}(E_n)}$ with respect to the orthonormal basis $\{n^{d/2}\chi_{Q_n(y)}\}_{y\in(\frac{1}{n}\mathbb{Z})^d\cap S_n}$. It is important to note that calculating the matrix elements of $\widetilde{H}_n$ requires knowledge only of $\{V_f\}_{f\in\Lambda_n}$ where we have $\Lambda_n := \left\{ f_y : y \in (n^{-1}\mathbb{Z})^d \cap S_n \right\}$ and $V_{f_y} = f_y(V) = V(y)$.

We have so far shown that the Assumption (i) holds, and we are left to show that the discretization we have chosen satisfies Assumption (ii). In particular, we need to demonstrate that our discretization satisfies (11.57). That is the topic of the following theorem.

Theorem 11.22. *Let $V \in C(\mathbb{R}^d)$ be sectorial as defined in* (4.5) *satisfying $|V(x)| \to \infty$ as $|x| \to \infty$, and let $h = -\Delta + V$ with $\mathcal{D}(h) = C_c^\infty(\mathbb{R}^d)$ and let $H = h^{**}$. Let H_n be as in* (11.66). *Then there exists z_0 such that $\|(H - z_0)^{-1} - (H_n - z_0)^{-1}E_n\| \to 0$, as $n \to \infty$.*

Remark 11.23 (Proof of Theorem 4.5). Note that we immediately have

$$\text{Theorem 11.22 } + \text{ Proposition 11.16 } \Rightarrow \text{ Theorem 4.5.}$$

Thus, the rest of the section is devoted to prove Theorem 11.22.

We shall treat the discretizations in a similar way as the continuous case, namely by "rotating" the operator into symmetric position with respect to the real axis and then, by taking the real part, we are dealing with a sequence of self-adjoint invertible operators. Before we prove this theorem we will need a couple of lemmas. We recall the following definition.

Definition 11.24 (Collectively compact). *A set $\mathcal{T} \subset B(\mathcal{H})$ is called collectively compact if the set $\{Tx : T \in \mathcal{T}, \|x\| \leq 1\}$ has compact closure.*

Lemma 11.25. *Let $\{K_n\}$ be a collectively compact operator sequence and $K_n^* \to 0$ strongly. Then $\|K_n\| \to 0$.*

Proof. It is well known that on any compact set $\mathcal{B}$ the strong convergence $K_n^* \to 0$ turns into norm convergence: $\sup\{\|K_n^* x\| : x \in \mathcal{B}\} \to_n 0$. Since $\mathcal{B} := \mathrm{cl}\{K_n x : \|x\| \leq 1, n \in \mathbb{N}\}$ is compact, we get

$$\|K_n\|^2 = \|K_n^* K_n\| = \sup\{\|K_n^* K_n x\| : \|x\| \leq 1\} \leq \sup\{\|K_n^* y\| : y \in \mathcal{B}\} \to 0 \quad \text{as} \quad n \to \infty.$$

□

We also need a modification of Lemma 11.18.

Lemma 11.26. *Let $\{A_n\}$ be m-sectorial with common semi-angle $\beta < \pi/2$ and denote $B_n = \mathrm{Re}\, A_n$. Assume that $\{E_n\}$ is a sequence of orthogonal projections, converging strongly to identity and such that $A_n E_n = E_n A_n E_n$ and $B_n E_n = E_n B_n E_n$. Assume further that $\{B_n^{-1}\}$ is uniformly bounded. If $\{B_n^{-1} E_n\}$ is collectively compact, then so is $\{A_n^{-1} E_n\}$.*

Proof. Denote by $B_n^{1/2}$ the unique self-adjoint non-negative square root of B_n. By [57, VI-Theorem 3.2 on p.337] for each A_n there exists a bounded symmetric operator C_n satisfying $\|C_n\| \leq \tan(\beta)$ and such that $A_n = B_n^{1/2}(1 + \mathrm{i}\, C_n) B_n^{1/2}$. Writing

$$A_n^{-1} = \int_0^\infty e^{-tA_n}\, dt$$

we conclude that $E_n A_n^{-1} E_n = A_n^{-1} E_n$ and likewise for B_n^{-1}. Assume now that $\{B_n^{-1} E_n\}$ is collectively compact. But then so is $\{(B_n + t)^{-1} E_n\} = \{B_n^{-1} E_n (I + t B_n^{-1})^{-1} E_n\}$ and writing, compare [57, V (3.43) on p.282],

$$B_n^{-1/2} E_n = \frac{1}{\pi} \int_0^\infty t^{-1/2} (B_n + t)^{-1} E_n\, dt$$

we see that $\{B_n^{-1/2} E_n\}$ is also collectively compact and $B_n^{-1/2} E_n = E_n B_n^{-1/2} E_n$. Finally $\{A_n^{-1} E_n\}$ is then collectively compact as well since $A_n^{-1} E_n$ is of the form $B_n^{-1/2} E_n T_n$ with T_n uniformly bounded. □

Proof of Theorem 11.22. Note that it is clear from the definition of H_n and the assumption on V that $\mathrm{Num}(H_n) \subset \{re^{i\rho} : -\theta_2 \leq \rho \leq \theta_1, r \geq 0\}$ for all n. Thus, since H_n is bounded and by Proposition 11.20 we can choose any point $z_0 \in \mathbb{C}$ such that z_0 has a positive distance d to the closed sector $\{re^{i\rho} : -\theta_2 \leq \rho \leq \theta_1, r \geq 0\}$, and both $R(H, z_0) = (H - z_0)^{-1}$ and $R(H_n, z_0) = (H_n - z_0)^{-1}$ for every n will exist. Moreover, $R(H_n, z_0)$ are uniformly bounded for all n, since for every x, $\|x\| = 1$,

$$\|(H_n - z_0)x\| \geq |\langle (H_n - z_0)x, x\rangle| \geq |\langle H_n x, x\rangle - z_0| \geq d.$$

Note that by Lemma 11.25 it suffices to show that (i) $R(H_n, z_0)^* E_n \to R(H, z_0)^*$ strongly, and (ii) $\{R(H_n, z_0) E_n - R(H, z_0)\}$ is collectively compact.

To see (i) observe that $C_c^\infty(\mathbb{R}^d)$ is a common core for H and for H_n. Hence by [57, VIII-Theorem 1.5 on p.429], the strong resolvent convergence $R(H_n, z_0)^* \to R(H, z_0)^*$ will follow if we show that

$H_n^*\psi \to H^*\psi$ as $n \to \infty$ for any $\psi \in C_c^\infty(\mathbb{R}^d)$. Then the strong convergence $R(H_n, z_0)^* E_n \to R(H, z_0)^*$ follows as well. Note that

$$\|H_n^*\psi - H^*\psi\| \leq \left\| \sum_{j=1}^d |\Phi_n|^2(P_j) E_n \psi - \sum_{j=1}^d P_j^2 \psi \right\| + \|(\overline{V}_n - \overline{V})\psi\|. \tag{11.67}$$

Also, $|\Phi_n|^2(P_j) = n(\tau_{-1/ne_j} - I)n(\tau_{1/ne_j} - I)$, where $\tau_z\psi(x) = \psi(x-z)$ and $\{e_j\}$ is the canonical basis for $\mathbb{R}^d$. Moreover, for $\psi \in C_c^\infty(\mathbb{R}^d)$,

$$E_n\psi = \sum_{y \in (\frac{1}{n}\mathbb{Z})^d \cap S_n} (\Psi_n * \psi)(y)\chi_{Q_n(y)}, \quad \Psi_n = \rho_n \otimes \ldots \otimes \rho_n, \quad \rho_n = n\chi_{[-\frac{1}{2n}, \frac{1}{2n})},$$

where S_n was defined in (11.65). Thus, it follows from easy calculus manipulations and basic properties of convolution that $|\Phi_n|^2(P_j)E_n\psi = \sum_{y \in (\frac{1}{n}\mathbb{Z})^d} (\Psi_n * \tilde{\rho}_1 *_j \tilde{\rho}_2 *_j \psi'')(y)\chi_{Q_n(y)}$, where $\tilde{\rho}_1 = n\chi_{[-1/n,0]}$, $\tilde{\rho}_2 = n\chi_{[0,1/n]}$ and $*_j$ denotes the convolution operation in the jth variable. By standard properties of the convolution we have that $\Psi_n * \tilde{\rho}_1 *_j \tilde{\rho}_2 *_j \psi'' \to \psi''$ uniformly as $n \to \infty$. Thus, since $\psi \in C_c^\infty(\mathbb{R}^d)$, the first part of the right hand side of (11.67) tends to zero as $n \to \infty$. Due to the continuity of V and the bounded support of ψ it also follows easily that $\|(\overline{V}_n - \overline{V})\psi\| \to 0$ as $n \to \infty$.

To see (ii) we use the same trick as in the proof of Proposition 11.20. In particular, first set $z_0 = -e^{i\alpha}$ (which is clearly in the resolvent set of H_n for $\alpha = (\theta_1 - \theta_2)/2$) then let $A_n(\alpha) = e^{-i\alpha}(H_n - z_0)$ and further $H_n(\alpha) = \mathrm{Re}\, A_n(\alpha)$. Note that, by Lemma 11.26, we would be done if we could show that $\{H_n(\alpha)^{-1}\}$ is uniformly bounded and $\{H_n(\alpha)^{-1}E_n\}$ is collectively compact as that would yield collective compactness of $\{A_n(\alpha)^{-1}E_n\}$ and hence of $\{R(H_n, z_0)E_n\}$. To establish collective compactness, note that

$$H_n(\alpha) = \cos\alpha\, E_n \sum_{j=1}^d |\Phi_n|^2(P_j)E_n + \cos\vartheta(x)|V_n(x)| + 1, \tag{11.68}$$

where ϑ is defined in (11.62). Thus $\|H_n(\alpha)^{-1}\| \leq 1$ and by applying Lemma 11.27 we are now done. $\square$

Lemma 11.27. *Let $H_n(\alpha)$ be given by (11.68). Then the set $\{H_n(\alpha)^{-1}E_n\}$ is collectively compact.*

Proof. We shall show that if we choose an arbitrary sequence $\{\psi_n\} \subset L^2(\mathbb{R}^d)$ satisfying $\|\psi_n\| \leq 1$, then the sequence $\{\varphi_n\}$ where $\varphi_n = H_n(\alpha)^{-1}E_n\psi_n$, is relatively compact in $L^2(\mathbb{R}^d)$. The compactness argument is based on the Rellich's criterion.

Lemma 11.28 (Rellich's criterion ([78] Theorem XIII.65)). *Let $F(x)$ and $G(\omega)$ be two measurable nonnegative functions becoming larger than any constant for all large enough $|x|$ and $|\omega|$. Then*

$$S = \{\varphi \ : \ \int |\varphi(x)|^2 dx \leq 1, \ \int F(x)|\varphi(x)|^2 dx \leq 1, \ \int G(\omega)|\mathcal{F}\varphi(\omega)|^2 d\omega \leq 1\}$$

is a compact subset of $L^2(\mathbb{R}^d)$.

To prove Lemma 11.27 we proceed as follows. First we conclude that $\{\varphi_n\}$ is a bounded sequence itself. Then, in order to be able to define suitable functions F, G we need to approximate the sequence by another one of the form $\Psi_n * \varphi_n$. This approximation shall satisfy $\lim_{n\to\infty} \|\Psi_n * \varphi_n - \varphi_n\| = 0$ and this is very similar to the standard result on local uniform convergence of mollifications of continuous functions. Then the Rellich's criterion holds for $\Psi_n * \varphi_n$ with $F(x)$ essentially given by $|V(x)|$ and $G(\omega)$ by $|\omega|^2$. We then conclude that the sequence $\{\Psi_n * \varphi_n\}$ is relatively compact. But since $\lim_{n\to\infty} \|\Psi_n * \varphi_n - \varphi_n\| = 0$, the sequence $\{\varphi_n\}$ is relatively compact as well, completing the argument.

More precisely, since $|\vartheta(x)| \leq \alpha < \pi/2$ we have from (11.68)

$$|\langle H_n(\alpha)\varphi_n, \varphi_n\rangle| \geq \cos\alpha \left(\langle \sum_{j=1}^d |\Phi_n|^2(P_j)\varphi_n, \varphi_n\rangle + \langle |V_n|\varphi_n, \varphi_n\rangle \right) + \|\varphi_n\|^2. \tag{11.69}$$

But $|\langle H_n(\alpha)\varphi_n,\varphi_n\rangle|$ is bounded not only from below but also from above. Indeed, $|\langle H_n(\alpha)\varphi_n,\varphi_n\rangle| = |\langle E_n\psi_n,\varphi_n\rangle| \leq \|H_n(\alpha)^{-1}E_n\|\|\psi_n\|^2$. Thus, we conclude first from (11.69) that the sequence $\{\varphi_n\}$ is bounded. Next, in view of (11.69), there exist constants $C_1, C_2 > 0$ such that for all $n \in \mathbb{N}$

$$\langle\sum_{j=1}^{d}|\Phi_n|^2(P_j)\varphi_n,\varphi_n\rangle \leq C_1, \tag{11.70a}$$

$$\langle|V_n|\varphi_n,\varphi_n\rangle \leq C_2. \tag{11.70b}$$

First we use the bound (11.70a). Letting $\mathcal{F}$ denote the Fourier transform , we have that $(\mathcal{F}\Phi_n(P_j)\varphi_n)(\omega) = \Phi_n(\omega_j)(\mathcal{F}\varphi_n)(\omega)$, for a.e. ω and for $1 \leq j \leq d$. Letting $\Theta_n(\omega) = \frac{\sin(\omega/2n)}{\omega/2n}$, an application of the Fourier transform to (11.70a) along with Plancherel's theorem yield

$$\int_{\mathbb{R}^d}|(\mathcal{F}\varphi_n)(\omega)|^2\sum_{1\leq j\leq d}|\omega_j\Theta_n(\omega_j)|^2\,d\omega \leq C_1.$$

Moreover, since $|\Theta_n(\omega)| \leq 1$ for all ω, we get

$$\int_{\mathbb{R}^d}|\omega|^2|\Theta_n(\omega_1)\cdots\Theta_n(\omega_d)|^2|(\mathcal{F}\varphi_n)(\omega)|^2\,d\omega \leq C_1. \tag{11.71}$$

We now define the approximation $\Psi_n*\varphi_n$. Let $\Psi_1(z) = \chi_{[-1/2,1/2]^d}(z)$ and further $\Psi_n(z) = n^d\Psi_1(nz)$, where $\chi_A(z)$ is the usual characteristic function for the set A. We shall prove below that $\lim_{n\to\infty}\|\Psi_n * \varphi_n - \varphi_n\| = 0$, which in particular shows that the sequence $\{\Psi_n * \varphi_n\}$ is bounded. Observe then that $(\mathcal{F}\Psi_n)(\omega) = \Theta_n(\omega_1)\cdots\Theta_n(\omega_d)$. Therefore we obtain from (11.71)

$$\int_{\mathbb{R}^d}|\omega|^2|\mathcal{F}(\Psi_n*\varphi_n)(\omega)|^2 d\omega \leq C_1,$$

which shows that we can choose $G(\omega)$ to be (a constant times) $|\omega|^2$.

We still need to establish the growth function $F(x)$ for $\Psi_n * \varphi_n$. Consider φ_n. It is of the form $\varphi_n = (E_n + E_nB_nE_n)^{-1}E_n\psi_n$ and hence $E_n\varphi_n = \varphi_n$. Therefore φ_n vanishes outside S_n and we can essentially replace V_n by V in the inequality (11.70b). To that end, put $F(x) = \min_{|y|\geq|x|}|V(y)|$. Then with some constant C_3

$$\int_{\mathbb{R}^d}F(x)|(\Psi_n*\varphi_n)(x)|^2dx \leq C_3. \tag{11.72}$$

In view of the bounds (11.71), (11.72) and since the sequence $\{\Psi_n * \varphi_n\}_{n\in\mathbb{N}}$ is bounded in L^2, Rellich's criterion implies that $\{\Psi_n * \varphi_n\}_{n\in\mathbb{N}}$ is a relatively compact sequence and it therefore follows that $\{\varphi_n\}_{n\in\mathbb{N}}$ is relatively compact, thus finishing the proof. Hence, our only remaining obligation is to show that $\lim_{n\to\infty}\|\Psi_n * \varphi_n - \varphi_n\| = 0$. This result is very similar to the standard result on local uniform convergence of mollifications of continuous functions.

Let $z \in \mathbb{R}^d$ and define the shift operator τ_z on $L^2(\mathbb{R}^d)$ by $\tau_z f(x) = f(x - z)$. Now observe that by Minkowski's inequality for integrals it follows that

(11.73)

$$\|\Psi_n*\varphi_n - \varphi_n\| \leq \int_{\mathbb{R}^d}\|\tau_{\frac{1}{n}z}\varphi_n - \varphi_n\||\Psi_1(z)|\,dz = \int_{[-1/2,1/2]^d}\|e^{i\frac{z_d}{n}P_d}\ldots e^{i\frac{z_1}{n}P_1}\varphi_n - \varphi_n\|\,dz.$$

The claim follows from an ϵ/d argument and (11.73) combined with the dominated convergence theorem (recall that $\{\varphi_n\}$ is bounded): we need to show that for fixed $z \in [-1/2,1/2]^d$ and for any $1 < j \leq d$,

$$\lim_{n\to\infty}\left\|e^{i\frac{z_j}{n}P_j}\ldots e^{i\frac{z_1}{n}P_1}\varphi_n - e^{i\frac{z_{j-1}}{n}P_{j-1}}\ldots e^{i\frac{z_1}{n}P_1}\varphi_n\right\| = 0, \quad \lim_{n\to\infty}\left\|e^{i\frac{z_1}{n}P_1}\varphi_n - \varphi_n\right\| = 0. \tag{11.74}$$

Since $e^{i\frac{z_j}{n}P_j}e^{i\frac{z_k}{n}P_k} = e^{i\frac{z_k}{n}P_k}e^{i\frac{z_j}{n}P_j}$ and $\|e^{i\frac{z_j}{n}P_j}\cdots e^{i\frac{z_1}{n}P_1}\| \leq 1$ for $1 \leq j,k \leq d$, (11.74) will follow if we can show that $\|(e^{i\frac{z_j}{n}P_j} - I)\varphi_n\| \to 0$ as $n \to \infty$. Note that, by the choice of the projections E_n, it follows that for $1 \leq j \leq d$, $|((e^{i\frac{z_j}{n}P_j} - I)\varphi_n)(x)| \leq |((e^{i\frac{1}{n}P_j} - I)\varphi_n)(x)|$, for $0 \leq z_j \leq 1/2$ and

$x \in \mathbb{R}^d$. Also, $|((e^{i\frac{z_j}{n}P_j} - I)\varphi_n)(x)| \leq |((e^{-i\frac{1}{n}P_j} - I)\varphi_n)(x)|$ for $-1/2 \leq z_j < 0$. However the bound $\sum_{1\leq j\leq d} \|\Phi_n(P_j)\varphi_n\|^2 \leq C_1$ implies that $\lim_{n\to\infty} \|(e^{\pm i\frac{1}{n}P_j} - I)\varphi_n)\| = 0$, which proves the claim. □

12. Proofs of theorems in section 5

Proof of Theorem 5.2. **Step I:** Clearly, $\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} \geq \mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{G}} \geq \mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{G}}$, and $\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} \geq \mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{A}} \geq \mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{G}}$. We start by showing that $\mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{G}} \geq 2$. For $n, m \in \mathbb{N} \setminus \{1\}$ let

$$B_{n,m} := \begin{pmatrix} 1/m & & & & 1 \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ 1 & & & & 1/m \end{pmatrix} \in \mathbb{C}^{n\times n}$$

and for a sequence $\{l_n\}_{n\in\mathbb{N}} \subset \mathbb{N} \setminus \{1\}$ set

$$A := \bigoplus_{n=1}^{\infty} B_{l_n, n+1}.$$

Clearly, A defines an invertible operator on $l^2(\mathbb{N})$. Furthermore, we define $b = \{b_j\} \in l^2(\mathbb{N})$ such that

$$b_j = \begin{cases} \dfrac{1}{n+2} & j = \sum_{i=1}^{n} l_i + 1, \quad n \in \mathbb{N}_0 \\ 0 & \text{otherwise.} \end{cases}$$

Let also $C_m := \mathrm{diag}\{1/m, 1, 1, \ldots\}$ and note that its inverse is given by $\mathrm{diag}\{m, 1, 1, \ldots\}$. We argue by contradiction and suppose that there is a General tower of algorithms Γ_n of height one such that $\Gamma_n(A, b) \to \Xi(A, b)$ as $n \to \infty$ for arbitrary A and b. For every A, b and $k \in \mathbb{N}$ let $N(A, b, k)$ denote the smallest integer such that the evaluations from $\Lambda_{\Gamma_k}(A, b)$ only take the matrix entries $A_{ij} = \langle Ae_j, e_i\rangle$ with $i, j \leq N(A, b, k)$ and the entries b_i with $i \leq N(A, b, k)$ into account. To obtain a particular counterexample (A, b) we construct sequences $\{l_n\}_{n\in\mathbb{N}}$ and $\{k_n\}_{n\in\mathbb{Z}_+}$ inductively such that A and b are given by $\{l_n\}$ as above but $\Gamma_{k_n}(A, b) \nrightarrow \Xi(A, b)$. As a start, set $k_0 = l_0 := 1$. The sequence $\{x_j^{(1)}\}_{j\in\mathbb{N}} := (C_2)^{-1}P_1 b$ has a 1 at its first entry and since, by assumption, $\Gamma_k \to \Xi$, there is a k_1 such that, for all $k \geq k_1$, the first entry of $\Gamma_k(C_2, P_1 b)$ is closer to 1 than $1/2$. Then, choose $l_1 > N(C_2, P_1 b, k_1) - l_0$. Now, for $n > 1$, suppose that $l_0, \ldots, l_{n-1}$ and $k_0, \ldots, k_{n-1}$ are already chosen. Set $s_n := \sum_{i=0}^{n-1} l_i$. Then also $P_{s_n} b$ is already determined and

$$x_{s_n}^{(n)} = 1, \quad \text{where} \quad \{x_j^{(n)}\}_{j\in\mathbb{N}} := (B_{l_1,2} \oplus B_{l_2,3} \oplus \ldots \oplus B_{l_{n-1},n} \oplus C_{n+1})^{-1} P_{s_n} b.$$

Since, by assumption, $\Gamma_k \to \Xi$, there is a k_n such that for all $k \geq k_n$

$$|x_{s_n}^{(n,k)} - 1| \leq 1/2, \quad \text{where} \quad \{x_j^{(n,k)}\}_{j\in\mathbb{N}} := \Gamma_k(B_{l_1,2} \oplus B_{l_2,3} \oplus \ldots \oplus B_{l_{n-1},n} \oplus C_{n+1}, P_{s_n} b).$$

Now, choose $l_n > N(B_{l_1,2} \oplus B_{l_2,3} \oplus \ldots \oplus B_{l_{n-1},n} \oplus C_{n+1}, P_{s_n} b, k_n) - l_0 - l_1 - \ldots - l_{n-1}$. By this construction we get for the resulting A and b that for every n

$$\Gamma_{k_n}(A, b) = \Gamma_{k_n}(B_{l_1,2} \oplus B_{l_2,3} \oplus \ldots \oplus B_{l_{n-1},n} \oplus C_{n+1}, P_{s_n} b).$$

In particular $\lim_{k\to\infty} \Gamma_k(A, b)$ does not exist in $l^2(\mathbb{N})$, a contradiction.

Step II: To show that $\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} \leq 2$, let A be invertible and $Ax = b$ with the unknown x. Since P_m are compact projections converging strongly to the identity, we get that the ranks $\mathrm{rk}\, P_m = \mathrm{rk}(AP_m) = \mathrm{rk}(P_n A P_m)$ for every m and all $n \geq n_0$ with an n_0 depending on m and A. Then, obviously, $P_m A^* P_n A P_m$ is an invertible operator on $\mathrm{Ran}(P_m)$, and we can define

$$\Gamma_{m,n}(A, b) := \begin{cases} \{0\}_{j\in\mathbb{N}} & \text{if } \sigma_1(P_m A^* P_n A P_m) \leq \frac{1}{m} \\ (P_m A^* P_n A P_m)^{-1} P_m A^* P_n b & \text{otherwise.} \end{cases}$$

Note that for every A, b, m, n in view of Proposition 10.1 and any standard algorithm for finite dimensional linear problems, these approximate solutions can be computed by finitely many arithmetic operations on finitely many entries of A and b, hence $\Gamma_{m,n}$ are general algorithms in the sense of Definition 2.3 and require only a finite number of arithmetic operations. Moreover, they converge to $y_m := (P_m A^* A P_m)^{-1} P_m A^* b$ as $n \to \infty$. It is well known that y_m is also a (least squares) solution of the optimization problem $\|AP_m y - b\| \to \min$, that is

$$\|AP_m y_m - b\| \le \|AP_m x - b\| \le \|A\|\|P_m x - A^{-1}b\| = \|A\|\|P_m x - x\| \to 0$$

as $m \to \infty$. Therefore $\|y_m - x\| = \|P_m y_m - x\|$ is not greater than

$$\|A^{-1}\|\|A(P_m y_m - x)\| = \|A^{-1}\|\|AP_m y_m - b\| \le \|A^{-1}\|\|A\|\|P_m x - x\| \to 0,$$

which yields the convergence $y_m \to x$ and finishes the proof of Step II.

Step III: Let f be a bound on the dispersion of A. The smallest singular values of the operators AP_m are uniformly bounded below by $\|A^{-1}\|^{-1}$ which, together with $\|P_{f(m)} AP_m - AP_m\| \to 0$, yields that the limes inferior of the smallest singular values of $P_{f(m)} AP_m$ is positive, hence the inverses of the operators $B_m := P_m A^* A P_m$ and $C_m := P_m A^* P_{f(m)} A P_m$ on the range of P_m exist for sufficiently large m and have uniformly bounded norm. Moreover, $\|B_m^{-1} - C_m^{-1}\| \le \|B_m^{-1}\|\|C_m - B_m\|\|C_m^{-1}\|$ tend to zero as $m \to \infty$.

This particularly implies that the norms $\|y_m - (P_m A^* P_{f(m)} A P_m)^{-1} P_m A^* b\|$ with y_m as above tend to zero as $m \to \infty$, and we easily conclude that the norms $\|y_m - \Gamma_{m,f(m)}(A,b)\|$ tend to zero as well. With the convergence $\|y_m - x\| \to 0$ from the previous proof, now also $\|x - \Gamma_{m,f(m)}(A,b)\| \to 0$ holds as $m \to \infty$, which is the assertion $\mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{A}} \le 1$. Clearly $\mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{G}} \ge 1$. □

Remark 12.1. The technique used with uneven sections to obtain the bound $\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} \le 2$ is also referred to as asymptotic Moore-Penrose inversion as well as modified (or non-symmetric) finite section method in the literature, although written in a different form, and is widely used (see e.g. [45, 46, 53, 84, 91]). Also the idea to exploit bounds on the off diagonal decay is considered e.g. in [44] or in the theory of band-dominated operators and operators of Wiener type (cf. [63, 77, 83]).

Proof of Theorem 5.3. We proceed in a similar way as in the previous proofs. Let $\{l_n\}_{n\in\mathbb{N}}$ be some sequence of integers $l_n \ge 2$. Define

$$A := \bigoplus_{n=1}^{\infty} B_{l_n} - I, \qquad B_n := \begin{pmatrix} 1 & & & & 1 \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ 1 & & & & 1 \end{pmatrix} \in \mathbb{C}^{n\times n}.$$

Clearly, such A are invertible and their inverses have norm one. Suppose that $\{\Gamma_k\}$ is a height-one General tower of algorithms which in its kth step only reads information contained in the first $N(A,k) \times N(A,k)$ entries of the input A. In order to find a counterexample we again construct an appropriate sequence $\{l_n\} \subset \mathbb{N} \setminus \{1\}$ by induction: For $C := \operatorname{diag}\{1, 0, 0, 0, \ldots\}$ one obviously has $\|(C-I)^{-1}\| = \infty$. As a start, choose $k_0 := 1$ and $l_1 > N(C - I, k_0)$. Now, suppose that $l_1, \ldots, l_n$ are already chosen. Then the operator given by the matrix $B_{l_1} \oplus \ldots \oplus B_{l_n} \oplus C - I$ is not invertible, hence there exists a k_n such that, for every $k \ge k_n$,

$$\Gamma_k(B_{l_1} \oplus \ldots \oplus B_{l_n} \oplus C - I) > 2.$$

Now finish the construction by choosing $l_{n+1} > N(B_{l_1} \oplus \ldots \oplus B_{l_n} \oplus C - I, k_n) - l_1 - l_2 - \ldots - l_n$.

So, we see that

$$\Gamma_{k_n}(A) = \Gamma_{k_n}(B_{l_1} \oplus \ldots \oplus B_{l_n} \oplus C - I) \nrightarrow \|A^{-1}\| = 1, \quad n \to \infty,$$

a contradiction. Thus $\mathrm{SCI}(\Xi,\Omega_1)_{\mathrm{G}} \geq 2$. In order to prove the equalities over Ω_1 and Ω_2 we introduce the numbers

$$\begin{aligned}
\gamma &:= \|A^{-1}\|^{-1} = \min\{\sigma_1(A), \sigma_1(A^*)\}\\
\gamma_m &:= \min\{\sigma_1(AP_m), \sigma_1(A^*P_m)\}\\
\gamma_{m,n} &:= \min\{\sigma_1(P_nAP_m), \sigma_1(P_nA^*P_m)\}\\
\delta_{m,n} &:= \min\{k/m : k \in \mathbb{N}, k/m \geq \sigma_1(P_nAP_m) \text{ or } k/m \geq \sigma_1(P_nA^*P_m)\}
\end{aligned}$$

and note that $\gamma_m \downarrow_m \gamma$, and $\gamma_{m,n} \uparrow_n \gamma_m$ for every fixed m. Moreover, $\{\delta_{m,n}\}_n$ is bounded and monotone, and $\gamma_{m,n} \leq \delta_{m,n} \leq \gamma_{m,n} + 1/m$. Thus, $\{\delta_{m,n}\}_n$ converges for every m, and for $\epsilon > 0$ there is an m_0, and for every $m \geq m_0$ there is an $n_0 = n_0(m)$ such that

$$|\gamma - \delta_{m,n}| \leq |\gamma - \gamma_m| + |\gamma_m - \gamma_{m,n}| + |\gamma_{m,n} - \delta_{m,n}| \leq \epsilon/3 + \epsilon/3 + 1/m \leq \epsilon \tag{12.1}$$

whenever $m \geq m_0$ and $n \geq n_0(m)$. Since $\delta_{m,n}$ and hence $\Gamma_{m,n}(A) := \delta_{m,n}^{-1}$ can again be computed with finitely many arithmetic operations by Proposition 10.1 this provides an Arithmetic tower of algorithms of height two, hence easily completes the proof for Ω_1 and Ω_2. On Ω_3 we apply (12.1) with $n = f(m)$ and straightforwardly check that $\Gamma_m(A) := \delta_{m,f(m)}^{-1}$ does the job. □

13. Proofs of theorems in section 6

Proof of Theorems 6.1, 6.2 and 6.4. We start with Theorem 6.2. Suppose that there exists a sequence $\{\tilde{\Gamma}_n\}$ of general algorithm $\tilde{\Gamma}_n : \Omega \to \mathbb{N}^k$ such that for any $A \in \Omega$, $d(\Gamma_{\tilde{\Gamma}_n(A)_k,\ldots,\tilde{\Gamma}_n(A)_1}(A), \Xi(A)) < \frac{1}{n}$. Define, for $n \in \mathbb{N}$, the mapping $\hat{\Gamma}_n : \Omega \to \mathcal{M}$ by

$$\hat{\Gamma}_n(A) = \Gamma_{\tilde{\Gamma}_n(A)_k,\ldots,\tilde{\Gamma}_n(A)_1}(A), \tag{13.1}$$

and $\Lambda_{\hat{\Gamma}_n}(A) = \Lambda_{\Gamma_{\tilde{\Gamma}_n(A)_k,\ldots,\tilde{\Gamma}_n(A)_1}}(A) \cup \Lambda_{\tilde{\Gamma}_n}(A)$. If we can show that $\hat{\Gamma}_n$ is a general algorithm the proof is finished, as that would, via (13.1), imply that there is a height one tower of algorithms for the computational problem $\{\Xi, \Omega, \mathcal{M}, \Lambda\}$, and that violates the assumption that $\mathrm{SCI}(\Xi, \Omega, \mathcal{M}, \Lambda)_{\mathrm{G}} \geq 2$.

To prove that $\hat{\Gamma}_n$ is a general algorithm, we first note that (i) and (ii) in Definition 2.3 are immediately satisfied, and thus we only concentrate on showing (iii). Suppose that $B \in \Omega$ and $f(A) = f(B)$ for all $f \in \Lambda_{\hat{\Gamma}_n}(A)$. We claim that $\Lambda_{\hat{\Gamma}_n}(B) = \Lambda_{\hat{\Gamma}_n}(A)$ and observe that this would imply (iii). Indeed, the assumption implies that $f(A) = f(B)$ for all $f \in \Lambda_{\tilde{\Gamma}_n}(A)$ so, by (iii) in Definition 2.3 we have that $\Lambda_{\tilde{\Gamma}_n}(B) = \Lambda_{\tilde{\Gamma}_n}(A)$, moreover, it follows that $\tilde{\Gamma}_n(A) = \tilde{\Gamma}_n(B)$. Now, again by using (iii) of Definition 2.3, since for all $f \in \Lambda_{\Gamma_{\tilde{\Gamma}_n(A)_k,\ldots,\tilde{\Gamma}_n(A)_1}}(A)$ we have $f(A) = f(B)$, it follows that $\Lambda_{\Gamma_{\tilde{\Gamma}_n(A)_k,\ldots,\tilde{\Gamma}_n(A)_1}}(B) = \Lambda_{\Gamma_{\tilde{\Gamma}_n(A)_k,\ldots,\tilde{\Gamma}_n(A)_1}}(A)$. Thus, since $\tilde{\Gamma}_n(A) = \tilde{\Gamma}_n(B)$ this implies the claim and we are done.

Note that the proof of Theorem 6.4 is almost identical, thus we omit the details. For Theorem 6.1 just notice that indeed, if such integers $n_k(m), \ldots, n_1(m)$ exist then $\tilde{\Gamma}_m(A) = \Gamma_{n_k(m),\ldots,n_1(m)}(A)$ would define a General tower of height one.

□

14. Proofs of theorems in Section 7

Proof of Theorem 7.1. **Step I:** The assertion for Ξ_1 is easy, so we start by discussing Ξ_2. To see that $\mathrm{SCI}(\Xi_2, \Omega_1)_{\mathrm{G}} > 1$, we argue by contradiction and assume that there is a General tower of algorithms $\{\Gamma_n\}_{n\in\mathbb{N}}$ of height one which can answer this question. Then there exist strictly increasing sequences $\{n_k\}_{k\in\mathbb{N}}, \{i_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ such that, for the sequence $\{a_i\}_{i\in\mathbb{N}}$ which has 1s exactly at the positions i_k, the algorithms Γ_{n_k}, applied to $\{a_i\}_{i\in\mathbb{N}}$, answer $\Gamma_{n_k}(\{a_i\}) =$ *No* for all n_k. This is proved by induction: Set $\{a_i^1\}_{i\in\mathbb{N}} := \{0, 0, \ldots\}$. Then there is an n_1 such that $\Gamma_n(\{a_i^1\}) =$ *No* for all $n \geq n_1$. Further, the set of evaluations $\Lambda_{\Gamma_{n_1}}(\{a_i^1\})$ of Γ_{n_1} is finite, i.e. in simple words Γ_{n_1} only looks at a finite number of entries, lets say entries with index less than i_1. Next, assume that i_k, n_k are already chosen for $k = 1, \ldots, m$. Let

$\{a_i^{m+1}\}_{i\in\mathbb{N}}$ denote the sequence which has entry 1 exactly at the positions $i_1,\dots,i_m$ and zeros everywhere else. Then, by the assumption that $\Gamma_n(\{a_i^{m+1}\}) \to \Xi_2(\{a_i^{m+1}\}) = \textit{No}$ as $n\to\infty$, there is an n_{m+1} greater than n_m such that $\Gamma_n(\{a_i^{m+1}\}) = \textit{No}$ for all $n \geq n_{m+1}$. Since $\Gamma_{n_{m+1}}$ evaluates only finitely many entries of $\{a_i^{m+1}\}_{i\in\mathbb{N}}$ there exists an $i_{m+1} > i_m$ such that $\Gamma_{n_{m+1}}$ only looks at the positions less than i_{m+1} of $\{a_i^{m+1}\}$.

Now consider the sequence $\{a_i\}_{i\in\mathbb{N}}$ which has entry 1 exactly at the positions i_k, $k\in\mathbb{N}$. Then we observe that $\Gamma_{n_k}(\{a_i\}) = \Gamma_{n_k}(\{a_i^k\}) = \textit{No}$ for every k, taking (iii) in Definition 2.3 into account, hence $\lim_k \Gamma_{n_k}(\{a_i\}) = \textit{No} \neq \Xi_2(\{a_i\})$, a contradiction.

To see that $\mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{A}} \leq 2$ define $\Gamma_{m,n}(\{a_i\}_{i\in\mathbb{N}}) = \textit{Yes}$ when $\sum_{i=1}^n a_i > m$. Here $\Lambda_{\Gamma_{m,n}}(\{a_i\})$ consists of the evaluations $f_j : \{a_i\}_{i\in\mathbb{N}} \mapsto a_j$ with $j=1,\dots,n$. Obviously, these mappings $\Gamma_{m,n}$ are General algorithms in the sense of Definition 2.3. If we define $\Gamma_m(\{a_i\}) = \textit{Yes}$ when $\sum_{i=1}^\infty a_i > m$, then we can observe that $\Gamma_{m,n}(\{a_i\}) \to \Gamma_m(\{a_i\})$ as $n\to\infty$ and $\Gamma_m(\{a_i\}) \to \Xi_2(\{a_i\})$ as $m\to\infty$. Thus, $\mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{G}} \leq \mathrm{SCI}(\Xi_2,\Omega_1)_{\mathrm{A}} \leq 2$, and we have shown the assertions for Ξ_2.

Step II: Note that it is easy to translate Ξ_3 and Ξ_4 into the previous problems: just take an enumeration of the elements of $\mathbb{N}^2$, that is a bijection $\mathbb{N}\to\mathbb{N}^2$, $k\mapsto(i(k),j(k))$ in order to regard $\{a_{i,j}\}$ as the sequence $\{a_{i(k),j(k)}\}_k$, which yields that Ξ_3 (Ξ_4) is equivalent to Ξ_1 (Ξ_2, respectively). Hence, $\mathrm{SCI}(\Xi_3,\Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi_3,\Omega_2)_{\mathrm{A}} = 1$ and $\mathrm{SCI}(\Xi_4,\Omega_2)_{\mathrm{G}} = \mathrm{SCI}(\Xi_4,\Omega_2)_{\mathrm{A}} = 2$. ☐

Before we continue with the proofs we need to introduce some helpful background. Equip the set Ω_1 of all sequences $\{x_i\}_{i\in\mathbb{N}} \subset \{0,1\}$ with the following metric:

(14.1)
$$d_{\mathrm{B}}(\{x_i\},\{y_i\}) := \sum_{n\in\mathbb{N}} 3^{-n}|x_n - y_n|.$$

The resulting metric space is known as the Cantor space. By the usual enumeration of the elements of $\mathbb{N}^2$ this metric translates to a metric on the set Ω_2 of all matrices $A=\{a_{i,j}\}_{i,j\in\mathbb{N}}$ with entries in $\{0,1\}$. Similarly, we do this for the set Ω_3 of all matrices $A=\{a_{i,j}\}_{i,j\in\mathbb{Z}}$ with entries in $\{0,1\}$. In each case this gives a complete metric space, hence a so called Baire space, i.e. it is of second category (in itself). To make this precise we recall the following definitions:

Definition 14.1 (Meager set). *A set $S\subset\Omega$ in a metric space Ω is nowhere dense if every open set $U\subset\Omega$ has an open subset $V\subset U$ such that $V\cap S=\emptyset$, i.e. if the interior of the closure of S is empty. A set $S\subset\Omega$ is meager (or of first category) if it is an at most countable union of nowhere dense sets. Otherwise S is nonmeager (or of second category).*

Notice that every subset of a meager set is meager, as is every countable union of meager sets. By the Baire category theorem, every (nonempty) complete metric space is nonmeager.

Definition 14.2 (Initial segment). *We call a finite matrix $\sigma\in\mathbb{C}^{n\times m}$ an initial segment for an infinite matrix $A\in\Omega_2$ and say that A is an extension of σ if σ is in the upper left corner of A. In particular, $\sigma = P_nAP_m$ for some $n,m\in\mathbb{N}$, where we, with slight abuse of notation, consider $P_nAP_m\in\mathbb{C}^{n\times m}$. P_n is as usual the projection onto $\mathrm{span}\{e_j\}_{j=1}^n$, where $\{e_j\}_{j\in\mathbb{N}}$ is the canonical basis for $l^2(\mathbb{N})$.*

Similarly, a finite matrix $\sigma\in\mathbb{C}^{(2n+1)\times(2m+1)}$ is an initial segment for an infinite matrix $B\in\Omega_3$ if σ is in the center of B i.e. $\sigma=\tilde{P}_nB\tilde{P}_m$ where $\tilde{P}_n$ is the projection onto $\mathrm{span}\{e_j\}_{j=-n}^n$, where $\{e_j\}_{j\in\mathbb{Z}}$ is the canonical basis for $l^2(\mathbb{Z})$. We denote that A is an extension of σ by $\sigma\subset A$, and the set of all extensions of σ by $E(\sigma)$.

The notion of extension extends in an obvious way to finite matrices.

Notice that the set $E(\sigma)$ of all extensions of σ is a nonempty open and closed neighborhood for every extension of σ.

Lemma 14.3. *Let $\{\Gamma_n\}_{n\in\mathbb{N}}$ be a sequence of General algorithms mapping $\Omega_2 \to \mathcal{M}$, $T \subset \Omega_2$ be a nonempty closed set, and $S \subset T$ be a nonmeager set (in T) such that $\xi = \lim_{n\to\infty} \Gamma_n(A)$ exists and is the same for all $A \in S$. Then there exists an initial segment σ and a number n_0 such that $E^T(\sigma) := T \cap E(\sigma)$ is not empty, and such that $\Gamma_n(A) = \xi$ for all $A \in E^T(\sigma)$ and all $n \geq n_0$. The same statement is true if we consider Ω_3 instead of Ω_2.*

Proof. We are in a complete metric space T. Since $S = \bigcup_{k\in\mathbb{N}} S_k$ with $S_k := \{A \in S : \Gamma_n(A) = \xi \ \forall n \geq k\}$ and S is nonmeager, not all of the S_k can be meager, hence there is a nonmeager S_k, and we set $n_0 := k$. Now, let A be in the closure $\overline{S_{n_0}}$, i.e. there is a sequence $\{A_j\} \subset S_{n_0}$ converging to A. Note that by assumption (i) in Definition 2.3 and the fact that Γ_n are General algorithms, we have that, for every fixed $n \geq n_0$, $|\Lambda_{\Gamma_n}(A)| < \infty$. Thus, by (ii) in Definition 2.3, the General algorithm Γ_n only depends on a finite part of A, in particular $\{A_f\}_{f\in\Lambda_{\Gamma_n}(A)}$ where $A_f = f(A)$. Since each $f \in \Lambda_{\Gamma_n}(A)$ represents a coordinate evaluation of A and by the definition of the metric d_B in (14.1), it follows that for all sufficiently large j, $f(A) = f(A_j)$ for all $f \in \Lambda_{\Gamma_n}(A)$. By assumption (iii) in Definition 2.3, it then follows that $\Lambda_{\Gamma_n}(A_j) = \Lambda_{\Gamma_n}(A)$ for all sufficiently large j. Hence, by assumption (ii) in Definition 2.3, we have that $\Gamma_n(A) = \Gamma_n(A_j) = \xi$ for all sufficiently large j. Thus, $\Gamma_n(A) = \xi$ for all $n \geq n_0$ and all $A \in \overline{S_{n_0}}$. Since S_{n_0} is not nowhere dense, we can choose a point $\tilde{A}$ in the interior of $\overline{S_{n_0}}$ and fix a sufficiently large initial segment σ of $\tilde{A}$ such that $E^T(\sigma)$ is a subset of $\overline{S_{n_0}}$. The assertion of the lemma now follows. The extension of the proof to Ω_3 is clear. $\square$

Roughly speaking, this shows that there is a nice open and closed nonmeager subspace of T for which $\lim_{n\to\infty} \Gamma_n(A)$ exists even in a uniform manner. Note that this result particularly applies to the case $T = \Omega$.

Proof of Theorem 7.2. **Step I:** $\mathrm{SCI}(\Xi_5, \Omega_2)_{\mathrm{G}} \geq 3$. We argue by contradiction and assume that there is a height two tower $\{\Gamma_r\}$, $\{\Gamma_{r,s}\}$ for Ξ_5, where Γ_r denote, as usual, the pointwise limits $\lim_{s\to\infty} \Gamma_{r,s}$. We will inductively construct initial segments $\{\sigma_n\}$ with $\sigma_{n+1} \supset \sigma_n$ yielding an infinite matrix $A \supset \sigma_n$ for all $n \in \mathbb{N}$, such that $\lim_{r\to\infty} \Gamma_r(A)$ does not exist. We construct $\{\sigma_n\}$ with the help of two sequences of subsets $\{T_n\}$ and $\{S_n\}$ of Ω, with the properties that $T_{n+1} \subset T_n$, each T_n is closed, and either $T_n = \Omega$ or there is an initial segment $\sigma \in \mathbb{C}^{m\times m}$ where $m \geq n$ such that T_n is the set of all extensions of σ with all the remaining entries in the first n columns being zero.

Suppose that we have chosen T_n. Note that the subset of all matrices in T_n with one particular entry being fixed is closed in T_n, hence the set of all matrices with one particular column being fixed is closed (as an intersection of closed sets). The latter set has no interior points in T_n, hence is nowhere dense in T_n. This provides that the set of all matrices in T_n for which a particular column has only finitely many 1s is a countable union of nowhere dense sets in T_n, hence is meager in T_n. Taking intersections we find that the set of all matrices in T_n having only finitely many 1s in each of their columns is meager in T_n as well. Let R be its complement in T_n, i.e. the nonmeager set of all matrices $A \in T_n$ with $\Xi_5(A) =$ *Yes*.

Clearly, $R = \bigcup_{r\in\mathbb{N}} R_r$ with $R_r := \{A \in R : \Gamma_k(A) = \textit{Yes}\ \forall k \geq r\}$, and there is an r_n such that $S_n := R_{r_n}$ is nonmeager in T_n. Note that $\Gamma_{r_n,s}$ are General algorithms and $\Gamma_{r_n}(A) = \lim_{s\to\infty} \Gamma_{r_n,s}(A) =$ *Yes* for all $A \in S_n$. Thus, Lemma 14.3 applies and yields an initial segment σ_n, such that

$$E^{T_n}(\sigma_n) \neq \emptyset \quad \text{and } \Gamma_{r_n}(A) = \textit{Yes} \text{ for all } A \in E^{T_n}(\sigma_n). \tag{14.2}$$

Now, let $T_{n+1} \subset T_n$ be the (closed) set of all matrices in $E^{T_n}(\sigma_n)$ with all remaining [3] entries in the first $n+1$ columns being zero. Letting $T_0 = \Omega$ we have completed the construction.

The nested initial segments $\sigma_{n+1} \supset \sigma_n$ obviously yield a matrix $A \in \cap_{n=0}^{\infty} T_n$ and this A has only finitely many 1s in each of its columns. However, by the construction of $\{T_n\}$, we have that $A \in E^{T_n}(\sigma_n)$ for all $n \in \mathbb{N}$. Thus, $\Xi_5(A) =$ *No*, but by (14.2), $\Gamma_k(A) =$ *Yes* for infinitely many k.

[3] I.e. outside the initial segment σ_n.

Step II: $\mathrm{SCI}(\Xi_5,\Omega_2)_{\mathrm{A}} \leq 3$. Let $\Gamma_{k,m,n}(\{a_{i,j}\}_{i,j\in\mathbb{N}}) = $ *Yes* when $\sum_{i=1}^n a_{i,j} > m$ for at least one $j\in\{1,\dots,k\}$, and *No* otherwise. Here $\Lambda_{\Gamma_{k,m,n}}(\{a_{i,j}\})$ consists of the evaluations $f_{r,s} : \{a_{i,j}\}_{i,j\in\mathbb{N}} \mapsto a_{r,s}$ with $r = 1,\dots,n$ and $r = 1,\dots,k$. Further, define $\Gamma_{k,m}(\{a_{i,j}\}) = $ *Yes* when $\sum_{i=1}^\infty a_{i,j} > m$ for at least one $j\in\{1,\dots,k\}$, and let $\Gamma_k(\{a_{i,j}\}) = $ *Yes* when $\sum_{i=1}^\infty a_{i,j} = \infty$ for at least one $j\in\{1,\dots,k\}$. It is straightforwardly checked that this yields an Arithmetic tower of algorithms, hence $\mathrm{SCI}(\Xi_5,\Omega_2)_{\mathrm{A}} \leq 3$, and thus we have shown the assertions for Ξ_5.

Step III: $\mathrm{SCI}(\Xi_5,\Omega_2)_{\mathrm{G}} \leq \mathrm{SCI}(\Xi_6,\Omega_2)_{\mathrm{G}}$ and $\mathrm{SCI}(\Xi_6,\Omega_2)_{\mathrm{A}} \leq 4$. To see the former, suppose we are given $\{a_{i,j}\}$, then define a new matrix by $b_{i,j} := \max\{a_{i,s} : s = 1,\dots,j\}$. Then $\{a_{i,j}\}$ has a column with infinitely many non-zero entries if and only if $\{b_{i,j}\}$ has infinitely many columns with infinitely many non-zero entries. On the other hand, there is an Arithmetic tower of algorithms of height 4 for Ξ_6. Indeed, let $\Gamma_{l,k,m,n}(\{a_{i,j}\}) = $ *Yes* when $\sum_{i=1}^n a_{i,j} > m$ for more than l numbers $j\in\{1,\dots,k\}$. Also define $\Gamma_{l,k,m}(\{a_{i,j}\}) = $ *Yes* when $\sum_{i=1}^\infty a_{i,j} > m$ for more than l numbers $j\in\{1,\dots,k\}$, as well as $\Gamma_{l,k}(\{a_{i,j}\}) = $ *Yes* when $\sum_{i=1}^\infty a_{i,j} = \infty$ for more than l numbers $j\in\{1,\dots,k\}$. And finally set $\Gamma_l(\{a_{i,j}\}) = $ *Yes* when $\sum_{i=1}^\infty a_{i,j} = \infty$ for more than l numbers $j\in\mathbb{N}$. It is easy to see that this is an Arithmetic tower of algorithms, hence we have shown that $\mathrm{SCI}(\Xi_6,\Omega_2)_{\mathrm{A}} \leq 4$.

Step IV: $\mathrm{SCI}(\Xi_7,\Omega_2)_{\mathrm{G}} \geq 3$. Indeed, one can proceed as in the proof of Step I: The set of all matrices with infinitely many "finite columns" is meager since it is a subset of the union of the meager sets V_k of all matrices with the kth column being finite. Thus, the set $\{A : \Xi_7(A) = \textit{Yes}\}$ is nonmeager.

Then exactly the same construction as above yields a sequence of closed spaces T_n and a matrix A in their intersection which has only finitely many 1s in each of its columns, but $\Gamma_{r_n}(A) = $ *Yes* for all r_n of a certain sequence $\{r_n\}$, a contradiction.

Step V: $\mathrm{SCI}(\Xi_8,\Omega_3)_{\mathrm{G}} \geq 3$. The proof is very similar to the proof of Step I. In particular, we argue by contradiction and assume that there is a height two tower $\{\Gamma_r\}$, $\{\Gamma_{r,s}\}$ for Ξ_8. As above, we inductively construct initial segments $\{\sigma_n\}$ with $\sigma_{n+1} \supset \sigma_n$ yielding an infinite matrix $A \supset \sigma_n$ for all $n\in\mathbb{N}$, such that $\lim_{r\to\infty}\Gamma_r(A)$ does not exist. We construct $\{\sigma_n\}$ with the help of two sequences of subsets $\{T_n\}$ and $\{S_n\}$ of Ω, with the properties that $T_{n+1}\subset T_n$, each T_n is closed, and either $T_n = \Omega_3$ or there is an initial segment $\sigma\in\mathbb{C}^{(2m+1)\times(2m+1)}$ where $m\geq n$ such that T_n is the set of all extensions of σ with all $\pm n$th semi-columns being filled by n additional 1s and infinitely many 0s, and and all the other kth columns, $|k|\leq n-1$, are being filled with zeros. In particular, if $\{a_{i,j}\}_{i,j\in\mathbb{Z}}\in T_n$ then

$$\begin{aligned}\{a_{i,\pm n}\}_{i\in\mathbb{Z}} &= \{\dots,0,\underbrace{1,\dots,1}_{n\text{ times}},\sigma_{-m,\pm n},\dots,\sigma_{m,\pm n},\underbrace{1,\dots,1}_{n\text{ times}},0,\dots\}^T,\\ \{a_{i,k}\}_{i\in\mathbb{Z}} &= \{\dots,0,\sigma_{-m,k},\dots,\sigma_{m,k},0,\dots\}^T,\quad k\in\mathbb{Z}_+,\ |k|\leq n-1.\end{aligned} \tag{14.3}$$

Suppose that we have chosen T_n. We argue as in Step I and deduce that for $k\in\mathbb{Z}$ the set of all matrices in T_n with one of the two kth semi-columns being fixed is nowhere dense in T_n, hence the set of all matrices in T_n with (one of the two) kth semi-columns having finitely many 1s is meager in T_n. We conclude that the set of all matrices in T_n with one semi-column having finitely many 1s is meager, thus its complement in T_n, the set of all matrices with all semi-columns having infinitely many 1s, is nonmeager. Therefore the same holds for the superset $\{A\in T_n : \Xi_8(A) = \textit{Yes}\}$. Denoting this set by R we obviously have $R = \bigcup_{r\in\mathbb{N}} R_r$ with $R_r := \{A\in R : \Gamma_k(A) = \textit{Yes}\ \forall k\geq r\}$, and there is an r_n such that $S_n := R_{r_n}$ is nonmeager in T_n. Note that $\Gamma_{r_n,s}$ are General algorithms and $\Gamma_{r_n}(A) = \lim_{s\to\infty}\Gamma_{r_n,s}(A) = \textit{Yes}$ for all $A\in S_n$. Thus, Lemma 14.3 applies and yields an initial segment σ_n, such that

$$E^{T_n}(\sigma_n)\neq\emptyset\quad\text{and } \Gamma_{r_n}(A) = \textit{Yes} \text{ for all } A\in E^{T_n}(\sigma_n). \tag{14.4}$$

Now, let $T_{n+1}\subset T_n$ be the (closed) set of all matrices $\{a_{i,j}\}_{i,j\in\mathbb{N}}$ in $E^{T_n}(\sigma_n)$ with the property that (14.3) holds with $\sigma = \sigma_n$. Letting $T_0 = \Omega_3$ concludes the construction. The nested sequence $\{\sigma_n\}$again defines a matrix $A\in\cap_{n=0}^\infty T_n$ with the property that A has finitely many but at least k non-zero entries in the each

of its kth semi-column which gives $\Xi_8(A) = \textit{No}$, but, by (14.4), $\Gamma_k(A) = \textit{Yes}$ for infinitely many k, a contradiction.

Step VII: $\mathrm{SCI}(\Xi_7, \Omega_2)_{\mathrm{A}} \leq 3$ and $\mathrm{SCI}(\Xi_8, \Omega_3)_{\mathrm{A}} \leq 3$. This can again be proved by defining an appropriate tower of height 3 directly, as was done for Ξ_5 in Step II, for example. For Ξ_7

$$\Gamma_{k,m,n}(\{a_{i,j}\}_{i,j\in\mathbb{N}}) = \textit{Yes} \quad \Leftrightarrow \quad |\{j = 1, \ldots, m : \sum_{i=1}^{n} a_{i,j} < m\}| < k$$

may do the job. A more elegant way is to translate and to employ the already proved results on spectra, see Remark 14.4. □

With the lower bounds of the SCI of the decision problems Ξ_7 and Ξ_8 established we can now get the lower bounds of the SCI of spectra and essential spectra of operators.

Proof of Step III in the proof of Theorem 3.8 in Section 10. First, note that without loss of generality we may identify $\Omega_1 = \mathcal{B}(l^2(\mathbb{N}))$ with $\Omega = \mathcal{B}(X)$, where $X = \oplus_1^\infty X_n = l^2(\mathbb{N}, X_n)$ is the space of square-summable sequences of elements in X_n, where $X_n = l^2(\mathbb{Z})$. For this just choose any enumeration (i.e. bijection) $\mathbb{N} \to \mathbb{N} \times \mathbb{Z}$ to get $l^2(\mathbb{N}) \sim l^2(\mathbb{N} \times \mathbb{Z}) \sim X$. Second, for the sake of simplicity in notation we identify a sequence $\{a_i\}_{i\in\mathbb{N}}$ with its extension $\{a_i^e\}_{i\in\mathbb{Z}}$ where $a_i^e := a_i$ for $i > 0$ and $a_i^e := 1$ for all $i \leq 0$.

Let $a = \{a_i\}_{i\in\mathbb{N}}$ be a sequence with $a_i \in \{0, 1\}$. Define the operator $B_a \in \mathcal{B}(l^2(\mathbb{Z}))$ as follows

$$B_a e_i := \begin{cases} e_l : & \text{if } a_i^e = 1, l = \max\{j : j < i, a_j^e = 1\} \\ e_i : & \text{otherwise}, \end{cases} \tag{14.5}$$

where $\{e_j\}_{j\in\mathbb{Z}}$ is the canonical basis for $l^2(\mathbb{Z})$. Hence, B_a acts as a shift on the basis elements $\{e_i : a_i^e = 1\}$, and as the identity on all other basis elements. Clearly, B_a has norm 1 and if $a = \{a_i\}$ has infinitely many 1s then B_a is invertible and its inverse (the "reverse shift", actually given by its transpose B_a^T) has norm 1 as well. On the other hand, if a has only finitely many 1s then B_a is Fredholm of index -1 (again just look at the regularizer B_a^T). Moreover, notice that evaluating a finite number of entries of B_a requires the evaluation of only a finite number of a_is. Now, given a matrix $\{a_{i,j}\} \in \Omega$ we take the operators $B_k = B_{\{a_{i,k}\}_{i\in\mathbb{N}}}$ arising from the kth column of $\{a_{i,j}\}$ via (14.5), respectively, and define the diagonal operator

$$C := \bigoplus_{k=1}^{\infty} B_k$$

on X. Obviously, its norm is 1. If only finitely many B_k are non-invertible (i.e. only finitely many columns of $\{a_{i,j}\}$ do not have infinitely many nonzero entries) then the diagonal operator defined from the regularizers B_k^T is a regularizer of norm 1 for C, hence $\mathrm{sp}_{\mathrm{ess}}(C)$ is contained in the unit circle in that case. Otherwise, if infinitely many B_k are non-invertible then C is not Fredholm, i.e. $0 \in \mathrm{sp}_{\mathrm{ess}}(C)$. Therefore we can conclude the following: If there was a tower $\{\Gamma_m\}$, $\{\Gamma_{m,n}\}$ of height two for the essential spectrum (of C or its respective Ω_1-counterpart), then we would get a tower of height two for Ξ_7 by

$$\begin{aligned} \tilde{\Gamma}_{m,n}(\{a_{i,j}\}) &:= \textit{Yes} \text{ if and only if } \Gamma_{m,n}(C) \cap B_{1/2}(0) = \emptyset, \\ \tilde{\Gamma}_{m}(\{a_{i,j}\}) &:= \textit{Yes} \text{ if and only if } \Gamma_{m}(C) \cap B_{1/2}(0) = \emptyset, \end{aligned} \tag{14.6}$$

contradicting the fact that the SCI of Ξ_7 is three, which is established in Theorem 7.2. □

Proof of Step III in the proof of Theorem 3.7. As above, first observe that, without loss of generality, we may identify $\Omega_1 = \mathcal{B}(l^2(\mathbb{N}))$ with $\Omega = \mathcal{B}(X)$, where $X = \oplus_{n=-\infty}^{\infty} X_n$ in the l^2-sense and where $X_n = l^2(\mathbb{Z})$. Second, we consider sequences $a = \{a_i\}_{i\in\mathbb{Z}}$ over $\mathbb{Z}$ with $a_i \in \{0, 1\}$, and define respective operators

$B_a \in \mathcal{B}(l^2(\mathbb{Z}))$ with matrix representation $B_a = \{b_{k,i}\}$ by

$$b_{k,i} := \begin{cases} 1: & k = i \text{ and } a_k = 0 \\ 1: & k < i \text{ and } a_k = a_i = 1 \text{ and } a_j = 0 \text{ for all } k < j < i \\ 0: & \text{otherwise.} \end{cases}$$

Then B_a is again a shift on a certain subset of basis elements and the identity on the other basis elements, hence we have the following possible spectra:

- $\mathrm{sp}(B_a) \subset \{0,1\}$ if $\{a_i\}$ has finitely many 1s.
- $\mathrm{sp}(B_a) = \mathbb{T}$, the unit circle, if there are infinitely many $i > 0$ with $a_i = 1$ and infinitely many $i < 0$ with $a_i = 1$ (we say $\{a_i\}$ is two-sided infinite).
- $\mathrm{sp}(B_a) = \mathbb{D}$, the unit disc, if $\{a_i\}$ has infinitely many 1s, but only finitely many for $i < 0$ or finitely many for $i > 0$ (we say $\{a_i\}$ is one-sided infinite in that case).

Next for a matrix $\{a_{i,j}\}_{i,j\in\mathbb{Z}}$ we define the operator

$$C := \bigoplus_{k=-\infty}^{\infty} B_k \tag{14.7}$$

on X, where $B_k = B_{\{a_{i,k}\}_{i\in\mathbb{Z}}}$ corresponds to the column $\{a_{i,k}\}_{i\in\mathbb{Z}}$ in the above sense. Concerning its spectrum we have $\bigcup_{k\in\mathbb{Z}} \mathrm{sp}(B_k) \subset \mathrm{sp}(C) \subset \mathbb{D}$ since $\|C\| = 1$. Clearly, if one of the columns is one-sided infinite then $\mathrm{sp}(C) = \mathbb{D}$. The same holds true if for every $k \in \mathbb{N}$ there is a finite column with at least k 1s. Otherwise (that is if there is a number D such that for every column it holds that it either has less than D 1s or is two-sided infinite) the spectrum $\mathrm{sp}(C)$ is a subset of $\{0\} \cup \mathbb{T}$. Therefore if we had a height two tower $\{\Gamma_m\}$, $\{\Gamma_{m,n}\}$ for the computation of the spectrum of C or its counterpart in Ω_1 then

$$\begin{aligned} \tilde{\Gamma}_{m,n}(\{a_{i,j}\}) &:= \textit{Yes} \text{ if and only if } \Gamma_{m,n}(C) \cap B_{1/4}(1/2) = \emptyset \\ \tilde{\Gamma}_{m}(\{a_{i,j}\}) &:= \textit{Yes} \text{ if and only if } \Gamma_{m}(C) \cap B_{1/4}(1/2) = \emptyset \end{aligned} \tag{14.8}$$

would provide a height two tower for Ξ_8, contradicting Theorem 7.2. □

Remark 14.4. Note that Step VII in the proof of Theorem 7.2 follows immediately from the fact that the SCI of spectra and essential spectra of arbitrary operators are bounded by three and by translating respective height three towers similarly to (14.6) and (14.8).

Proof of Theorem 7.3. **Step I:** $\mathrm{SCI}(\Xi,\Omega_1)_{\mathrm{G}} > 2$. Note that this results follows almost directly from the techniques used in the proof of Step III in the proof of Theorem 3.7 above. Indeed, we may define C as in (14.7), and assume that there is a height two tower $\{\Gamma_m\}$, $\{\Gamma_{m,n}\}$ for the computation of Ξ. We can now argue exactly as done in the proof of Step III in the proof of Theorem 3.7, and deduce that we can define a height two tower for Ξ_8 from Section 7.1 as follows:

$$\begin{aligned} \tilde{\Gamma}_{m,n}(\{a_{i,j}\}) &:= \textit{Yes} \text{ if and only if } \Gamma_{m,n}(C - \tfrac{1}{2}I) = \textit{No} \\ \tilde{\Gamma}_{m}(\{a_{i,j}\}) &:= \textit{Yes} \text{ if and only if } \Gamma_{m}(C - \tfrac{1}{2}I) = \textit{No}. \end{aligned}$$

This would of course contradict Theorem 7.2.

Step II: $\mathrm{SCI}(\Xi,\Omega_1)_{\mathrm{A}} \le 3$. Let $A \in \mathcal{B}(l^2(\mathbb{N}))$ and define the numbers

$$\begin{aligned} \gamma &:= \min\{\sigma_1(A), \sigma_1(A^*)\} \\ \gamma_m &:= \min\{\sigma_1(AP_m), \sigma_1(A^*P_m)\} \\ \gamma_{m,n} &:= \min\{\sigma_1(P_nAP_m), \sigma_1(P_nA^*P_m)\} \\ \delta_{m,n} &:= \min\{2^{-m}k : k \in \mathbb{N}, 2^{-m}k \ge \sigma_1(P_nAP_m) \text{ or } 2^{-m}k \ge \sigma_1(P_nA^*P_m)\}, \end{aligned}$$

where $\sigma_1(B) := \inf\{\|B\xi\| : \xi \in X, \|\xi\| = 1\}$ for $B \in \mathcal{B}(l^2(\mathbb{N}))$ is the smallest singular value of B and where the operators of the form P_nAP_m are regarded as operators/matrices in $\mathcal{B}(\mathrm{Ran}P_m, \mathrm{Ran}P_n)$, respectively. It is well known that A is invertible if and only if $\gamma > 0$. Further, note that $\gamma_m \downarrow_m \gamma$, and that $\gamma_{m,n} \uparrow_n \gamma_m$ for every fixed m. The sequences $\{\delta_{m,n}\}_n$ are bounded and monotonically non-decreasing, and $\gamma_{m,n} \leq \delta_{m,n} \leq \gamma_{m,n} + 2^{-m} \leq \gamma_m + 2^{-m}$. Thus, for $\epsilon > 0$ there is an m_0, and for every $m \geq m_0$ there is an $n_0 = n_0(m)$ such that

$$|\gamma - \delta_{m,n}| \leq |\gamma - \gamma_m| + |\gamma_m - \gamma_{m,n}| + |\gamma_{m,n} - \delta_{m,n}| \leq \epsilon/3 + \epsilon/3 + 2^{-m} \leq \epsilon \tag{14.9}$$

whenever $m \geq m_0$ and $n \geq n_0(m)$. So we see that the numbers $\delta_{m,n}$ converge monotonically from below for every m as $n \to \infty$, and the respective limits form a non-increasing sequence w.r.t. m, tending to γ. Moreover, each $\delta_{m,n}$ can be computed with finitely many arithmetic operations by Proposition 10.1. Thus, if we define $\Gamma_{k,m,n}(A) := (\delta_{m,n} < k^{-1})$, we arrive at an Arithmetic tower of algorithms of height three

$$\mathrm{sp}^0(A) = \lim_{k\to\infty} \lim_{m\to\infty} \lim_{n\to\infty} \Gamma_{k,m,n}(A)$$

for our problem, hence $\mathrm{SCI}(\Xi, \Omega_1)_{\mathrm{A}} \leq 3$.

Step III: $\mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{A}} \leq \mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{A}} \leq 2$. If one considers operators for which a bound f on their dispersion is known, then choosing $n = f(m)$ turns (14.9) into

$$|\gamma - \delta_{m,f(m)}| \leq |\gamma - \gamma_m| + |\gamma_m - \gamma_{m,f(m)}| + |\gamma_{m,f(m)} - \delta_{m,f(m)}| \leq \epsilon/3 + \epsilon/3 + 2^{-m} \leq \epsilon \tag{14.10}$$

for large m taking $|\sigma_1(BP_m) - \sigma_1(P_{f(m)}BP_m)| \leq \|(I - P_{f(m)})BP_m\|$ into account. Therefore, a natural first guess for General algorithms could be $\tilde{\Gamma}_{k,m}(A) := (\delta_{m,f(m)} < k^{-1})$. Unfortunately, although $\delta_{m,f(m)}$ converges to γ as $m \to \infty$ by (14.10), this is not monotone in general. Hence, it might be the case that $\gamma = k^{-1}$, but $\delta_{m,f(m)}$ oscillates around k^{-1} such that $\{\tilde{\Gamma}_{k,m}(A)\}_m$ may not converge. To overcome this drawback, we consider the following algorithms

$$\Gamma_{k,m}(A) := (|\{i = 1, \dots, m : (\delta_{i,f(i)} < k^{-1})\}| + |\{i = 1, \dots, m : (\delta_{i,f(i)} < (k+1)^{-1})\}| > m).$$

This converges in any case as $m \to \infty$. Indeed: If $\gamma > k^{-1}$ then $\Gamma_{k,m}(A) =$ *No* for sufficiently large m. If $\gamma < (k+1)^{-1}$ then $\Gamma_{k,m}(A) =$ *Yes* for sufficiently large m. If $\gamma = k^{-1}$ then $|\{i = 1, \dots, m : (\delta_{i,f(i)} < (k+1)^{-1})\}|$ are the same for all sufficiently large m and it follows that $\Gamma_{k,m}(A)$ converges. Analogously for $\gamma = (k+1)^{-1}$. Finally, if $(k+1)^{-1} < \gamma < k^{-1}$ then again one of these families is uniformly bounded and the other only misses a finite number of points, and we again get convergence. Now, it is clear that $\mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{A}} \leq 2$.

Step IV: $\mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{G}} \geq \mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{G}} \geq 2$. If we assume that there is a General height-one-tower of algorithms $\{\Gamma_n\}$ over Ω_3 then we can again construct counterexamples very easily: For a decreasing sequence $\{a_i\}$ of positive numbers we consider the diagonal operator $A := \mathrm{diag}\{a_i\}$. Clearly, 0 belongs to the spectrum of A if and only if the a_is tend to zero. As a start, set $\{a_i^1\} := \{1, 1, \dots\}$, choose n_1 such that $\Gamma_n(\mathrm{diag}\{a_i^1\}) =$ *No* for all $n \geq n_1$, and i_1 such that $\Gamma_{n_1}(\mathrm{diag}\{a_i^1\})$ does not see the diagonal entries a_i^1 with indices $i \geq i_1$. This is possible by (iii) in Definition 2.3. Then set $\{a_i^2\} := \{1, 1, \dots, 1, 1/2, 1/2, \dots\}$ with $1/2$s starting at the i_1th position. If $n_1, \dots, n_{k-1}$ and $i_1, \dots, i_{k-1}$ are already chosen then pick n_k such that $\Gamma_n(\mathrm{diag}\{a_i^k\}) =$ *No* for all $n \geq n_k$, and i_k such that $\Gamma_{n_k}(\mathrm{diag}\{a_i^k\})$ doesn't see the diagonal entries a_i^k with indices $i \geq i_k$, and modify $\{a_i^k\}$ to $\{a_i^{k+1}\} := \{1, \dots, 2^{-k}, 2^{-k}, \dots\}$ with 2^{-k}s starting at the i_kth position. Now, the contradiction is as in the previous proofs and we see that $\mathrm{SCI}(\Xi, \Omega_2)_{\mathrm{G}} \geq \mathrm{SCI}(\Xi, \Omega_3)_{\mathrm{G}} \geq 2$. □

Proof of Theorem 7.13. **Step I:** We show that if $\mathrm{SCI}(\Xi, \Omega)_{\mathrm{G}} \leq m$ then Ξ is Δ_{m+1}. Let $p = \lim_i p_i$. Then

$$p = \textit{true} \quad \Leftrightarrow \quad \forall n \exists k (k \geq n \wedge p_k) \quad \Leftrightarrow \quad \exists n \forall k (k \leq n \vee p_k).$$

Further, let $\varphi : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$, $k \mapsto (\varphi_1(k), \varphi_2(k))$ be a bijection which enumerates all pairs of natural numbers, and note that

$$\exists n \exists m (p_{n,m}) \Leftrightarrow \exists k (p(\varphi_1(k), \varphi_2(k))), \qquad \forall n \forall m (p_{n,m}) \Leftrightarrow \forall k (p(\varphi_1(k), \varphi_2(k))),$$

for any family $(p_{n,m})_{n,m\in\mathbb{N}} \subset \mathcal{M}$. Thus, every limit in a tower of height m can be converted alternately into an expression with two quantifiers ($\forall\exists$ or $\exists\forall$), and then $m-1$ doubles $\exists\exists$ or $\forall\forall$ can be replaced by single quantifiers. This easily gives the claim (i) of Theorem 7.13.

Step II: We show that if Ξ is Σ_m or Π_m then $\mathrm{SCI}(\Xi,\Omega)_{\mathrm{G}} \le m$. As a start let $(p_i) \subset \mathcal{M}$ be a sequence. Then

$$(\forall i(p_i)) = \textit{true} \quad \Leftrightarrow \quad \left(\lim_{n\to\infty} \bigwedge_{i=1}^{n} p_i\right) = \textit{true}, \qquad (\exists i(p_i)) = \textit{true} \quad \Leftrightarrow \quad \left(\lim_{n\to\infty} \bigvee_{i=1}^{n} p_i\right) = \textit{true}.$$

Furthermore, the conjunction (disjunction) of limits coincides with the limit of the elementwise conjunction (disjunction), hence

$$\forall n_m \exists n_{m-1} \cdots \forall n_1 \Gamma_{n_m,\cdots,n_1} = \lim_{k_m} \lim_{k_{m-1}} \cdots \lim_{k_1} \bigwedge_{i_m=1}^{k_m} \bigvee_{i_{m-1}=1}^{k_{m-1}} \cdots \bigwedge_{i_1=1}^{k_1} \Gamma_{i_m,i_{m-1},\cdots,i_1}$$

and similarly for any other possible alternating quantifier form. Since the $\Gamma_{n_m,\cdots,n_1}$ in the alternating quantifier form at the left hand side are General algorithms, the right hand side obviously yields a tower of algorithms of height m.

Step III: Let $m \in \mathbb{N}$ be the smallest number with Ξ being Δ_{m+1}. In the above steps we have already seen that $m \le \mathrm{SCI}(\Xi,\Omega)_{\mathrm{G}} \le m+1$, and we next prove the following: If

$$\Xi(y) = \exists i \forall j (g_0(i,j,y)) = \forall n \exists m (g_1(n,m,y))$$

then $\Xi(y) = \lim_{k\to\infty} g(k,y)$ with a function g being easily derivable from g_0, g_1. The following construction is adopted from [43, Proofs of Theorems 1 and 3]. Fix y and define a function $h_0 : \mathbb{N} \to \mathcal{M}$ recursively as follows:

$i(1) := 1$, $j(1) := 1$, $h_0(1) := g_0(i(1), j(1), y)$.
`If` $h_0(l) =$ `true`
`then:` $i(l+1) := i(l)$, $j(l+1) := j(l) + 1$
`else:` $i(l+1) := i(l) + 1$, $j(l+1) := 1$.
$l := l + 1$.
$h_0(l) := g_0(i(l), j(l), y)$.

We observe that, if $\Xi(y) =$ *true* then $h_0(l)$ converges as $l \to \infty$ with limit *true*. Otherwise, the limit does not exist or is *false*. The same construction applies to $\neg(\forall n \exists m (g_1(n,m,y))) = \exists n \forall m \neg(g_1(n,m,y))$ and yields a function h_1 which converges to *true* if and only if $\Xi(y) =$ *false*. Clearly, exactly one of the functions h_0, h_1 converges to *true*. Now we derive the desired g from h_0 and h_1 as follows:

$\alpha(1) = 0$.
`If` $h_{\alpha(k)}(k) =$ `true`
`then:` $\alpha(k+1) := \alpha(k)$
`else:` $\alpha(k+1) := 1 - \alpha(k)$.
$k := k + 1$.
`If` $\alpha(k) = 0$
`then:` $g(k,y) :=$ `true`
`else:` $g(k,y) :=$ `false`.

This provides $\Xi(y) = \lim_{k\to\infty} g(k,y)$.

Next, let g_0 and g_1 be of the form $g_s(i,j,y) = \lim_r f^s_{i,j,r}(y)$, $s \in \{0,1\}$. Fix y. Then for every pair (i,j) there is an $r(i,j)$ such that $f^s_{u,v,r}(y) = g_s(u,v,y)$ for all $u \le i$, $v \le j$, $s \in \{0,1\}$ and $r \ge r(i,j)$. Thus, g is also of the form $g(k,y) = \lim_r f_{k,r}(y)$ with $f_{k,r}$ being defined by the above procedure applied to the functions $(i,j,y) \mapsto f^s_{i,j,k}(y)$ instead of $g_s(i,j,y)$ ($s \in \{0,1\}$).

Now we are left with iterating this argument: If both functions g_s ($s \in \{0,1\}$) are of the form $g_s(i,j,y) = \lim_{k_{m-1}} \lim_{k_{m-2}} \cdots \lim_{k_1} f^s_{i,j,k_{m-1},\cdots,k_1}(y)$ with certain General algorithms $f^s_{i,j,k_{m-1},\cdots,k_1}$, then also g is

of the form

$$g(k,y) = \lim_{k_{m-1}} \lim_{k_{m-2}} \cdots \lim_{k_1} f_{k,k_{m-1},\cdots,k_1}(y)$$

with $f_{k,k_{m-1},\cdots,k_1}$ being defined by the same procedure as before applied to the functions $(i,j,y) \mapsto f^s_{i,j,k_{m-1},\cdots,k_1}(y)$ instead of $g_s(i,j,y)$ ($s \in \{0,1\}$). The resulting functions $y \mapsto f_{k,k_{m-1},\cdots,k_1}(y)$ are General algorithms for every k, since their evaluation requires only finitely many evaluations of the General algorithms $f^s_{i,j,k_{m-1},\cdots,k_1}$. □

ACKNOWLEDGEMENTS

ACH would like to thank Caroline Series for pointing out the connection between the results of Doyle and McMullen in [38] and the work in [49]. It was this connection that initiated the work leading to this paper.

ACH acknowledges support from a Royal Society University Research Fellowship as well as the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/L003457/1.

REFERENCES

[1] B. Adcock and A. C. Hansen. Generalized sampling and infinite-dimensional compressed sensing. *Found. Comp. Math.*, (to appear).

[2] B. Adcock, A. C. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier: A new theory for compressed sensing. *arXiv:1302.0561*, 2014.

[3] W. Arveson. Improper filtrations for C^*-algebras: spectra of unilateral tridiagonal operators. *Acta Sci. Math. (Szeged)*, 57(1-4):11–24, 1993.

[4] W. Arveson. C^*-algebras and numerical linear algebra. *J. Funct. Anal.*, 122(2):333–360, 1994.

[5] W. Arveson. The role of C^*-algebras in infinite-dimensional numerical linear algebra. In *C^*-algebras: 1943–1993 (San Antonio, TX, 1993)*, volume 167 of *Contemp. Math.*, pages 114–129. Amer. Math. Soc., Providence, RI, 1994.

[6] G. Beer. *Topologies on closed and closed convex sets*, volume 268 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1993.

[7] C. Beltran, Luis, and M. Pardo. Smale's 17th problem: average polynomial time to compute affine and projective solutions. *J. Amer. Math. Soc*, 22(2), 2009.

[8] C. M. Bender. Making sense of non-Hermitian Hamiltonians. *Rep. Progr. Phys.*, 70(6):947–1018, 2007.

[9] C. M. Bender, D. C. Brody, and H. F. Jones. Complex extension of quantum mechanics. *Phys. Rev. Lett.*, 89(27):270401, 4, 2002.

[10] G. Beylkin and M. Mohlenkamp. Numerical operator calculus in higher dimensions. *Proc. Natl. Acad. Sci. USA*, 99(16):10246–10251, 2002.

[11] L. Blum, F. Cucker, M. Shub, S. Smale, and R. M. Karp. *Complexity and real computation*. Springer, New York, Berlin, Heidelberg, 1998.

[12] L. Blum, M. Shub, and S. Smale. On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bull. Amer. Math. Soc. (N.S.)*, 21(1):1–46, 1989.

[13] M. Blümlinger and R. Tichy. Topological algebras of functions of bounded variation i. *manuscripta mathematica*, 65(2):245–255, 1989.

[14] A. Böttcher. Pseudospectra and singular values of large convolution operators. *J. Integral Equations Appl.*, 6(3):267–301, 1994.

[15] A. Böttcher, H. Brunner, A. Iserles, and S. P. Nørsett. On the singular values and eigenvalues of the Fox-Li and related operators. *New York J. Math.*, 16:539–561, 2010.

[16] A. Böttcher, A. V. Chithra, and M. N. N. Namboodiri. Approximation of approximation numbers by truncation. *Integral Equations Operator Theory*, 39(4):387–395, 2001.

[17] A. Böttcher and B. Silbermann. *Analysis of Toeplitz operators*. Springer Monographs in Mathematics. Springer, 1990.

[18] L. Boulton. Projection methods for discrete Schrödinger operators. *Proc. London Math. Soc. (3)*, 88(2):526–544, 2004.

[19] A. Bove, G. Fano, G. Turchetti, and A. G. Teolis. Functional integration, Padé approximants, and the Schrödinger equation. *Journal of Mathematical Physics*, 16(2), 1975.

[20] N. P. Brown. Quasi-diagonality and the finite section method. *Math. Comp.*, 76(257):339–360, 2007.

[21] H. Brunner, A. Iserles, and S. P. Nørsett. The computation of the spectra of highly oscillatory Fredholm integral operators. *J. Integral Equations Appl.*, 23(4):467–519, 2011.

[22] P. Bürgisser and F. Cucker. Exotic quantifiers, complexity classes, and complete problems. *Foundations of Computational Mathematics*, (2):135–170.

[23] P. Bürgisser and F. Cucker. On a problem posed by Steve Smale. *Ann. of Math. (2)*, 174(3):1785–1836, 2011.

[24] P. Bürgisser and F. Cucker. *Condition - The Geometry of Numerical Algorithms.*, volume 349 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2013.

[25] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52(2):489–509, 2006.

[26] P. Casazza and G. Kutyniok. *Finite Frames: Theory and Applications*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, 2012.

[27] J. W. Cooley. An Improved Eigenvalue Corrector Formula for Solving the Schrödinger Equation for Central Fields. *Mathematics of Computation*, 15(76):pp. 363–374, 1961.

[28] G. Criscuolo, E. Minicozzi, and G. Trautteur. Limiting recursion and the arithmetic hierarchy. *ITA*, 9(3):5–12, 1975.

[29] E. B. Davies. A hierarchical method for obtaining eigenvalue enclosures. *Math. Comp.*, 69(232):1435–1455, 2000.

[30] E. B. Davies and M. Plum. Spectral pollution. *IMA J. Numer. Anal.*, 24(3):417–438, 2004.

[31] E. B. Davies and E. Shargorodsky. Level sets of the resolvent norm of a linear operator revisited. *arXiv:1408.2354*, 2014.

[32] P. Deift, L. C. Li, and C. Tomei. Toda flows with infinitely many variables. *J. Funct. Anal.*, 64(3):358–402, 1985.

[33] N. Dencker, J. Sjöstrand, and M. Zworski. Pseudospectra of semiclassical (pseudo-) differential operators. *Comm. Pure Appl. Math.*, 57(3):384–415, 2004.

[34] N. Dershowitz and Y. Gurevich. A natural axiomatization of computability and proof of Church's Thesis. *Bull. Symbolic Logic*, 14(3):299–350, 2008.

[35] L. Dickson. *Algebraic Theories*. Dover Books. S. Dover Publications, 1959.

[36] T. Digernes, V. S. Varadarajan, and S. R. S. Varadhan. Finite approximations to quantum systems. *Rev. Math. Phys.*, 6(4):621–648, 1994.

[37] D. L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.

[38] P. Doyle and C. McMullen. Solving the quintic by iteration. *Acta Math.*, 163(3-4):151–180, 1989.

[39] N. Dunford and J. Schwartz. *Linear Operators: General theory*. Pure and applied mathematics. Interscience Publishers, 1958.

[40] I. I. Ekeland and T. Turnbull. *Infinite-dimensional optimization and convexity*. Chicago lectures in mathematics. University of Chicago Press, Chicago, 1983.

[41] B. Engquist and A. Majda. Absorbing boundary conditions for numerical simulation of waves. *Proceedings of the National Academy of Sciences*, 74(5):1765–1766, 1977.

[42] J. Globevnik. Norm-constant analytic functions and equivalent norms. *Illinois J. Math.*, 20(3):503–506, 1976.

[43] E. M. Gold. Limiting recursion. *J. Symbolic Logic*, 30:28–48, 1965.

[44] K. Gröchenig and A. Klotz. Norm-controlled inversion in smooth banach algebras, i. *J. London Math. Society*, 88(1):49–64, 2013.

[45] K. Gröchenig, Z. Rzeszotnik, and T. Strohmer. Convergence analysis of the finite section method and banach algebras of matrices. *Integral Equations and Operator Theory*, 67(2):183–202, 2010.

[46] R. Hagen, S. Roch, and B. Silbermann. *C^*-algebras and numerical analysis*, volume 236 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 2001.

[47] A. C. Hansen. On the approximation of spectra of linear operators on Hilbert spaces. *J. Funct. Anal.*, 254(8):2092–2126, 2008.

[48] A. C. Hansen. Infinite-dimensional numerical linear algebra: theory and applications. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 466(2124):3539–3559, 2010.

[49] A. C. Hansen. On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators. *J. Amer. Math. Soc.*, 24(1):81–124, 2011.

[50] A. C. Hansen, B. Jennings, and M. Seidel. On the Solvability Complexity Index and the foundations of computational mathematics. *In preparation*.

[51] N. Hatano and D. R. Nelson. Localization transitions in non-hermitian quantum mechanics. *Phys. Rev. Lett.*, 77(3):570–573, Jul 1996.

[52] N. Hatano and D. R. Nelson. Vortex pinning and non-hermitian quantum mechanics. *Phys. Rev. B*, 56(14):8651–8673, Oct 1997.

[53] G. Heinig and F. Hellinger. The finite section method for Moore-Penrose inversion of Toeplitz operators. *Integral Equations Operator Theory*, 19(4):419–446, 1994.

[54] J. Hubbard, D. Schleicher, and S. Sutherland. How to find all roots of complex polynomials by newton's method. *Inventiones Mathematicae*, 146:1–33, 2000.

[55] A. Iserles. *A First Course in the Numerical Analysis of Differential Equations*. Cambridge texts in applied mathematics. Cambridge University Press, 1996.

[56] P. Junghanns, G. Mastroianni, and M. Seidel. On the stability of collocation methods for Cauchy singular integral equations in weighted L^p spaces. *Math. Nachr.*, 283(1):58–84, 2010.

[57] T. Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[58] L. Kiepert. Auflösung der gleichungen fünften grades. *Journal für die reine und angewandte Mathematik*, (v. 87), 1879.

[59] R. B. King. *Beyond the quartic equation*. Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA, 2009. Reprint of the 1996 original.

[60] S. C. Kleene. Recursive predicates and quantifiers. *Trans. Amer. Math. Soc.*, 53:41–73, 1943.

[61] S. C. Kleene. Arithmetical predicates and function quantifiers. *Trans. Amer. Math. Soc.*, 79:312–340, 1955.

[62] J. Knight. The Kleene-Mostowski hierarchy and the Davis-Mostowski hierarchy. In *Andrzej Mostowski and Foundational Studies*. IOS Press, 2008.

[63] M. Lindner. *Infinite matrices and their finite sections*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2006. An introduction to the limit operator method.

[64] M. Marletta. Neumann-dirichlet maps and analysis of spectral pollution for non-self-adjoint elliptic pdes with real essential spectrum. *IMA Journal of Numerical Analysis*, 2009.

[65] M. Marletta and S. Naboko. The finite section method for dissipative operators. *Mathematika*, 60:415–443, 7 2014.

[66] H. Mascarenhas, P. A. Santos, and M. Seidel. Quasi-banded operators, convolutions with almost periodic or quasi-continuous data, and their approximations. *J. Math. Anal. Appl.*, 418(2):938–963, 2014.

[67] C. McMullen. Families of rational maps and iterative root-finding algorithms. *Ann. of Math. (2)*, 125(3):467–493, 1987.

[68] C. McMullen. Braiding of the attractor and the failure of iterative algorithms. *Invent. Math.*, 91(2):259–272, 1988.

[69] A. Mostowski. On definable sets of positive integers. *Fund. Math.*, 34:81–112, 1947.

[70] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[71] E. Novak, I. H. Sloan, and H. Woźniakowski. Tractability of Approximation for Weighted Korobov Spaces on Classical and Quantum Computers. *Foundations of Computational Mathematics*, 4(2):121–156, 2004.

[72] P. Odifreddi. *Classical Recursion Theory (Volume I)*. North–Holland Publishing Co., Amsterdam, 1989.

[73] V. Ozolins, R. Lai, R. Caflisch, and S. Osher. Compressed modes for variational problems in mathematics and physics. *Proceedings of the National Academy of Sciences*, 2013.

[74] E. W. Packel and J. F. Traub. Information-based complexity. *Nature*, 328:29–33, 1987.

[75] M. B. Pour-El and I. Richards. The eigenvalues of an effectively determined self-adjoint operator are computable, but the sequence of eigenvalues is not. *Advances in Mathematics*, 63(1):1 – 41, 1987.

[76] H. Putnam. Trial and error predicates and the solution to a problem of Mostowski. *J. Symbolic Logic*, 30:49–57, 1965.

[77] V. Rabinovich, S. Roch, and B. Silbermann. *Limit operators and their applications in operator theory*, volume 150 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2004.

[78] M. Reed and B. Simon. *Analysis of operators*. Methods of Modern Mathematical Physics. Academic Press, 1978.

[79] H. Rogers, Jr. *Theory of recursive functions and effective computability*. MIT Press, Cambridge, MA, USA, 1987.

[80] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, Nov. 1992.

[81] L. K. Schubert. Iterated limiting recursion and the program minimization problem. *J. Assoc. Comput. Mach.*, 21:436–445, 1974.

[82] M. Seidel. On (N,ϵ)-pseudospectra of operators on Banach spaces. *J. Funct. Anal.*, 262(11):4916–4927, 2012.

[83] M. Seidel. Fredholm theory for band-dominated and related operators: a survey. *Linear Algebra Appl.*, 445:373–394, 2014.

[84] M. Seidel and B. Silbermann. Finite sections of band-dominated operators: l^p-theory. *Complex Anal. Oper. Theory*, 2(4):683–699, 2008.

[85] M. Seidel and B. Silbermann. Finite sections of band-dominated operators – norms, condition numbers and pseudospectra. In *Operator theory, pseudo-differential equations, and mathematical physics*, volume 228 of *Oper. Theory Adv. Appl.*, pages 375–390. Birkhäuser/Springer Basel AG, Basel, 2013.

[86] E. Shargorodsky. On the level sets of the resolvent norm of a linear operator. *Bull. Lond. Math. Soc.*, 40(3):493–504, 2008.

[87] E. Shargorodsky. On the limit behaviour of second order relative spectra of self-adjoint operators. *J. Spectr. Theory*, 2013.

[88] J. R. Shoenfield. On degrees of unsolvability. *Ann. of Math. (2)*, 69:644–653, 1959.

[89] M. Shub and S. Smale. On the existence of generally convergent algorithms. *J. Complexity*, 2(1):2–11, 1986.

[90] M. Shub and S. Smale. Complexity of Bézout's theorem. I. Geometric aspects. *J. Amer. Math. Soc.*, 6(2):459–501, 1993.

[91] B. Silbermann. Modified finite sections for Toeplitz operators and their singular values. *SIAM J. Matrix Anal. Appl.*, 24(3):678–692 (electronic), 2003.

[92] J. Sjöstrand and M. Zworski. Asymptotic distribution of resonances for convex obstacles. *Acta Mathematica*, 183(2):191–253, 1999.

[93] S. Smale. The fundamental theorem of algebra and complexity theory. *Bull. Amer. Math. Soc. (N.S.)*, 4(1):1–36, 1981.

[94] S. Smale. Complexity theory and numerical analysis. In *Acta numerica, 1997*, volume 6 of *Acta Numer.*, pages 523–551. Cambridge Univ. Press, Cambridge, 1997.

[95] S. Smale. The work of Curtis T McMullen. In *Proceedings of the International Congress of Mathematicians I, Berlin*, Doc. Math. J. DMV, pages 127–132. 1998.

[96] J. Traub, G. Wasilkowski, and H. Woźniakowski. *Information-based Complexity*. Academic Press, 1988.

[97] L. N. Trefethen and M. Embree. *Spectra and pseudospectra*. Princeton University Press, Princeton, NJ, 2005. The behavior of nonnormal matrices and operators.

[98] A. M. Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proc. London Math. Soc.*, S2-42(1):230.

[99] A. Wigderson. P, NP and mathematics – a computational complexity perspective. *Proceedings of the International Congress of Mathematicians, Madrid*, I, 2006.

[100] X. Zheng and K. Weihrauch. The arithmetical hierarchy of real numbers. In *Mathematical Foundations of Computer Science 1999*, volume 1672 of *Lecture Notes in Computer Science*, pages 23–33. Springer Berlin Heidelberg, 1999.

[101] M. Zworski. Resonances in physics and geometry. *Notices Amer. Math. Soc.*, 46(3):319–328, 1999.

DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE LONDON
E-mail address: J.Ben-Artzi@imperial.ac.uk

DAMTP, UNIVERSITY OF CAMBRIDGE
E-mail address: A.Hansen@damtp.cam.ac.uk

DEPARTMENT OF MATHEMATICS AND SYSTEMS ANALYSIS, AALTO UNIVERSITY
E-mail address: olavi.nevanlinna@aalto.fi

WEST SAXON UNIVERSITY OF APPLIED SCIENCES, ZWICKAU
E-mail address: markus.seidel@fh-zwickau.de